

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 12:51:27 ; Search time 522.566 Seconds  
(without alignments)  
15555.198 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_1\_143  
Perfect score: 143  
Sequence: 1 ttgccaccctcctctctgcg.....aactccatcactaggggtaa 143

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_in.\*
- 3: gb\_env.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pr.\*
- 9: gb\_ro.\*
- 10: gb\_sts.\*
- 11: gb\_sy.\*
- 12: gb\_un.\*
- 13: gb\_vi.\*
- 14: gb\_htg.\*
- 15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	100.0	4718	6	BD242766
2	143	100.0	4718	6	AR562498
3	143	100.0	4718	6	AX753251
4	143	100.0	4718	13	AF063497
5	138.8	97.1	4683	13	AF028704
6	137.2	95.9	4683	6	BD242775
7	137.2	95.9	4683	6	AR562507
8	135.6	94.8	4721	6	CQ972062
9	135.6	94.8	4721	6	CS073592
10	135.6	94.8	4721	6	AX753246
11	135.6	94.8	4721	13	AF513851
12	119	83.2	4718	6	BD242766
13	119	83.2	4718	6	AR562498
14	119	83.2	4718	6	AX753251
15	119	83.2	4718	13	AF063497
16	113.2	79.2	4722	13	AF028705
17	113.2	79.2	4726	6	AX753253
18	113.2	79.2	4726	13	AVU48704

19	112.2	78.5	4722	13	AF028705
20	111.6	78.0	4721	6	CQ972062
21	111.6	78.0	4721	6	CS073592
22	111.6	78.0	4721	6	AX753246
23	111.6	78.0	4721	13	AF513851
24	107.4	75.1	145	6	A46401
25	107.4	75.1	145	13	AA2L1TR1
26	107.4	75.1	145	13	AA2RSPORI
27	107.4	75.1	165	6	AR034135
28	107.4	75.1	165	6	BD218219
29	107.4	75.1	165	6	I16806
30	107.4	75.1	165	6	AR223306
31	107.4	75.1	165	6	AR58389
32	107.4	75.1	165	6	AX106702
33	107.4	75.1	207	6	AX703496
34	107.4	75.1	272	6	BD218218
35	107.4	75.1	272	6	AR223305
36	107.4	75.1	272	6	AR58388
37	107.4	75.1	272	6	AX106701
38	107.4	75.1	2116	13	AA2LEPT
39	107.4	75.1	4675	6	BD094552
40	107.4	75.1	4675	6	AX135805
41	107.4	75.1	4675	6	AX286292
42	107.4	75.1	4675	6	AX753252
43	107.4	75.1	4675	13	AA2CG
44	107.4	75.1	4679	6	AX282480
45	107.4	75.1	4679	13	AF043303

ALIGNMENTS

RESULT 1  
BD242766  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BD242766 4718 bp DNA linear PAT 17-JUL-2003  
Adeno-associated virus serum type 1 nucleic acid sequence, vector  
and host cell containing the same.  
BD242766  
BD242766.1 GI:33052536  
JP 2002529098-A/1.  
unidentified  
unidentified  
unclassified.  
1 (bases 1 to 4718)  
Wilson, J.M. and Xiao, W.  
Adeno-associated virus serum type 1 nucleic acid sequence, vector  
and host cell containing the same  
Patent: JP 2002529098-A 1 10-SEP-2002;  
THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA  
OS AAV-1  
PN JP 2002529098-A/1  
PD 10-SEP-2002  
PP 02-NOV-1999 JP 2000581227  
PR 05-NOV-1998 US 60/107114  
PI JAMES M WILSON, WEIDONG XIAO  
PC C12N15/09, A61K31/711, A61K48/00, A61P43/00, C12N1/15, C12N1/19, PC  
C12N1/21,  
PC C12N5/10, C12P21/02//A61K35/76, C12N15/00, C12N5/00 CC  
Adeno-associated virus serum type 1 nucleic acid sequence, CC  
vector and host  
cell containing the same  
Location/Qualifiers  
FH Key (335) . (2206)  
FT CDS (2223) . (4430) .  
FT CDS Location/Qualifiers  
1 . 4718  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

Query Match 100.0%; Score 143; DB 6; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 2.5e-21;

```
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGCCACTCCCTCTCTGCGCGCTGCTCGCTCGGTGGGGCTGCGGACCAAAAGTCCGC 60
Db 1 TTGCCACTCCCTCTCTGCGCGCTGCTCGCTCGGTGGGGCTGCGGACCAAAAGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGCCCAACCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGCGCCCAACCGAGCGAGCGCGCAGAGAGGAGTG 120
QY 121 GGCAACTCCATCACTAGGGGTAA 143
Db 121 GGCAACTCCATCACTAGGGGTAA 143

RESULT 2
AR562498
LOCUS AR562498 4718 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6759237.
ACCESSION AR562498
VERSION AR562498.1 GI:53976564
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4718)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors
and host cells containing same
JOURNAL Patent: US 6759237-A 1 06-JUL-2004;
The Trustees of The University of Pennsylvania; Philadelphia, PA
FEATURES
source
1. .4718
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 143; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 2.5e-21;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGCCACTCCCTCTCTGCGCGCTGCTCGCTCGGTGGGGCTGCGGACCAAAAGTCCGC 60
Db 1 TTGCCACTCCCTCTCTGCGCGCTGCTCGCTCGGTGGGGCTGCGGACCAAAAGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGCCCAACCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGCGCCCAACCGAGCGAGCGCGCAGAGAGGAGTG 120
QY 121 GGCAACTCCATCACTAGGGGTAA 143
Db 121 GGCAACTCCATCACTAGGGGTAA 143

RESULT 3
AR753251
LOCUS AR753251 4718 bp DNA linear PAT 23-JUN-2003
DEFINITION Sequence 6 from Patent EP1310571.
ACCESSION AR753251
VERSION AR753251.1 GI:32166108
KEYWORDS
SOURCE Adeno-associated virus 1
ORGANISM Adeno-associated virus 1
REFERENCE 1
AUTHORS Gao,G., Wilson,J.M. and Alvira,M.
TITLE A method of detecting and/or identifying adeno-associated virus
(AAV) sequences and isolating novel sequences identified thereby
JOURNAL Patent: EP 1310571-A 6 14-MAY-2003;
The Trustees of The University of Pennsylvania (US)
FEATURES
Location/Qualifiers
source
1. .4718
/organism="Adeno-associated virus 1"
```

```
/mol_type="unassigned DNA"
/db_xref="taxon:85106"

ORIGIN

Query Match 100.0%; Score 143; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 2.5e-21;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGCCACTCCCTCTCTGCGCGCTGCTCGCTCGGTGGGGCTGCGGACCAAAAGTCCGC 60
Db 1 TTGCCACTCCCTCTCTGCGCGCTGCTCGCTCGGTGGGGCTGCGGACCAAAAGTCCGC 60
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGCCCAACCGAGCGAGCGCGCAGAGAGGAGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGCGCCCAACCGAGCGAGCGCGCAGAGAGGAGTG 120
QY 121 GGCAACTCCATCACTAGGGGTAA 143
Db 121 GGCAACTCCATCACTAGGGGTAA 143

RESULT 4
AF063497
LOCUS AF063497 4718 bp DNA linear VRL 27-APR-1999
DEFINITION Adeno-associated virus 1, complete genome.
ACCESSION AF063497
VERSION AF063497.1 GI:4689096
KEYWORDS
SOURCE Adeno-associated virus 1
ORGANISM Adeno-associated virus 1
REFERENCE 1 (bases 1 to 4718)
AUTHORS Xiao,W., Chirmule,N., Berta,S.C., McCullough,B., Gao,G. and
Wilson,J.M.
TITLE Gene therapy vectors based on adeno-associated virus type 1
JOURNAL J. Virol. 73 (5), 3994-4003 (1999)
PUBMED 10196295
REFERENCE 2 (bases 1 to 4718)
AUTHORS Xiao,W. and Wilson,J.M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601
Spruce Street, Philadelphia, PA 19104, USA
FEATURES
Location/Qualifiers
source
1. .4718
/organism="Adeno-associated virus 1"
/mol_type="genomic DNA"
/db_xref="taxon:85106"
/note="AAV1"
335..2206
/codon_start=1
/product="nonstructural protein"
/protein_id="AAD27758.1"
/db_xref="GI:4689098"
/translation="MPGYEIVIKVPSDLDEHLPGISDFVSWAEKEWELPPDSMD
LNLIQAPLTVAEKLQDFLVQMRVSKAPALRVQPEKGSYFHLHLVETTVGVS
MYLGRFLSQIRDLKLVQTIYRGIEPLFPNFAVTKTRNGAGGKNKVDSCYIPNYLLPK
TQPELQAWTMMSEYISACLNLAERKLVAQRLTHVTSQTBQNKENLNPNSDAPVRS
KTSARYMLVGLVDRGITSKQWIEQDQASYISFNAASRSQIKAAALNAGKIMLPG
TKSADPYLVGPAPPADIKNRIYRIELNGYEPYAGSVFLQWAKRFKRNITWLPF
PATGTGNTIAEIAHAHVPGVCMVNTNFPCDVKMVIWMESEKMTAKVYESAKA
ILGSKRVVDQCKSSAOIDPTPIVTNTNMCAVIDGNSTTFEHQPLQDRMFKPEL
TRLEHDFGKVTQBEVFFRFAQDHVTEVAHEFVVRKGGANKRPAPDADKSEKRA
CPSVADPSTSDAEGAPVDADRIYQNKSRHAGMLQMLFPCKTCERMNQNFNICPTHGT
RDCSCFFGVSESQPVVRKRTYRKLCAIHLHLLGRAPEIACSCADLVNVDLDCVSEQ"
2223..4433
/codon_start=1
/product="capsid protein"
/protein_id="AAD27757.1"
/db_xref="GI:4689097"
/translation="MAADGYLPDWLEDLNLSGIREWMDLKPQAPKPKANQKQDDGRG
LVLPGYKYLGPFGNLDKSGEYPVNAADAEHDKAYDQQLKAGDNPYLRYNHADAFOE
RLQEDTSFGNLRGAVFOAKKRVLEPLGLVEGAKTAPGKKRPVQSPOEPDSSSIG
KTGQOPAKKRLNFGQTGDSSEVPDPQLGERPATPAAVGPTTMASSGGGAPMADNNEGA
```

DGVGNAGNWHCDSTWLGDRVITSTRTWALPTYNHLYKQISSASTGASNDNHYFG  
STPMGYDFNFRPHCHFSRDRWOLRNNMGRPKRLNFKNIOVKQEVTTNDGVTTIA  
NNLSTVQVFSDEYOLPYVLGSAHQGLPPADVFMIPYGYLTNNNGSAVGRSS  
FYCLEVFPQMLRTGNFTFSYTFEEVPHSHSYAHSOSLDRLMNPIDOLYLYARTQ  
NOSGSAQNDLLFSRGSAGMSVQPKNMLPGCYRQORVSKTKTNNNSNFTWTGASK  
YNLNGRESIINFGTAMASHKDDKFFPMGSMVIFGKSAGASATLNDNMVITDDEEI  
TDGHPHPSLMGFGGLKPNPPQILIKNTVPANPAEFASATFASFIQYSTQVSV  
IEWELQKENSXWNPVQYTSNYAKSANVDFTVNNGLYTEPRPIGTRYLTRLPL"

ORIGIN

Query Match 100.0%; Score 143; DB 13; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 2.5e-21;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTGGTGGGCGCTGCGGACCAAGGTCCGC 60  
DB 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTGGTGGGCGCTGCGGACCAAGGTCCGC 60  
  
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGGCCACCGAGCGAGCGGCGGAGGAGGTG 120  
DB 61 AGACGGCAGAGCTCTGCTCTGCGCGGCCACCGAGCGGAGCGGCGGAGGAGGTG 120  
  
QY 121 GGCAACTCCATCTAGGGGTAA 143  
DB 121 GGCAACTCCATCTAGGGGTAA 143

RESULT 5  
AF028704/c  
LOCUS  
DEFINITION Adeno-associated virus 6, complete genome.  
ACCESSION AF028704  
VERSION AF028704.1 GI:2766605  
KEYWORDS Adeno-associated virus 6  
ORGANISM Adeno-associated virus 6  
VIRUSES: sDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  
REFERENCE Rutledge,E.A., Halbert,C.L. and Russell,D.W.  
AUTHORS Infectious clones and vectors derived from adeno-associated virus  
TITLE (AAV) serotypes other than AAV type 2  
JOURNAL J. Virol. 72 (1), 309-319 (1998)  
PUBLISHED 9420229  
REFERENCE 2 (bases 1 to 4683)  
AUTHORS Rutledge,E.A. and Russell,D.W.  
TITLE Direct Submission  
JOURNAL Submitted (24-SEP-1997) Hematology, Univ. of Washington, Box  
357720, Seattle, WA 98195, USA  
FEATURES  
source Location/Qualifiers  
1..4683  
/organism="Adeno-associated virus 6"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:68558"  
/note="AAV6"  
320..2191  
/codon\_start=1  
/evidence=not\_experimental  
/product="nonstructural protein Rep78"  
/protein\_id="AAB95449.1"  
/db\_xref="GI:2766606"

CDS

translation="MPGFYRIVIKVPSDLDELHGGISDFVNVAAEKWELPPDSMD  
LNLEQAPLTVAEKLQDFLVQHRVSKAPEALFFVQPEKESYFHLHLIVETGVKS  
MVLGRFSQRIDLVQTIYRGIETPLNPAFVKTNRGAGGNKVDECIYNLLPK  
TQPELQWAMNMEYISACLNLAERKLVADLTHVSQTEQENKLNPNSDAPVIRS  
KTSARYMELGVLDITSEKQWIOEDQASYISFNAASNSRSQIKAADNAGKIMAL  
TKSAPDVLGAPADPKTRIVRIELNGLYDPAAGSVFLGWAQKRFGRKRVTFI  
PATTGNTIAEAIAHAPVFGVCNVNENPFNDKVMVWEEGKMTAKVYESAKA  
ILGSKRVQDKCKSSAQIDPTVIVTSNTNCAVIDGNSITTEHQOPLQDRMFKPEL  
TRLEHDFGKTVQEVKFEFFRWAQDHVTEVAHEFYVRKGANKRPPADKSEPKRA  
RDCSECFPGVSESQPVVRKRYRKLCAIHLHLLGRAPEIACSDLVNVDLDCVSEQ"  
2208..4418

CDS

ORIGIN

Query Match 97.1%; Score 138.8; DB 13; Length 4683;  
Best Local Similarity 98.6%; Pred. No. 2e-20;  
Matches 140; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTGGTGGGCGCTGCGGACCAAGGTCCGC 60  
DB 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTGGTGGGCGCTGCGGACCAAGGTCCGC 4624  
  
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGGCCACCGAGCGGAGCGGCGGAGGAGGTG 120  
DB 61 AGACGGCAGAGCTCTGCTCTGCGCGGCCACCGAGCGGAGCGGCGGAGGAGGTG 4564  
  
QY 121 GGCAACTCCATCTAGGGGTAA 142  
DB 4563 GGCAACTCCATCTAGGGGTAA 4542

RESULT 6

BD242775/c  
LOCUS  
DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector  
and host cell containing the same.  
ACCESSION BD242775.1 GI:33052545  
VERSION BD242775.1  
KEYWORDS JP 2002529098-A/10.  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 4683)  
AUTHORS Wilson,J.M. and Xiao,W.  
TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector  
and host cell containing the same  
JOURNAL Patent: JP 2002529098-A 10 10-SEP-2002;  
THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA  
COMMENT OS aav-6  
PN JP 2002529098-A/10  
PD 10-SEP-2002  
PF 02-NOV-1999 JP 2000581227  
PR 05-NOV-1998 US 60/107114  
PI JAMES M WILSON, WEIDONG XIAO  
PC C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC  
C12N1/21,  
PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC  
Adeno-associated virus serum type 1 nucleic acid sequence, CC  
vector and host  
CC cell containing the same  
FH Key Location/Qualifiers  
FT source 1..4683  
FT Location/Qualifiers  
1..4683  
/organism="unidentified"  
/mol\_type="genomic DNA"

```
ORIGIN
Patent: WO 2004108922-A 7 16-DEC-2004;
The Trustees of The University of Pennsylvania (US)
/db_xref="taxon:32644"

JOURNAL
Patent: WO 2004108922-A 7 16-DEC-2004;
The Trustees of The University of Pennsylvania (US)
/db_xref="taxon:272636"

FEATURES
source
Location/Qualifiers
1..4721
/organism="Adeno-associated virus"
/mol_type="unassigned DNA"
/db_xref="taxon:272636"

ORIGIN
Query Match 94.8%; Score 135.6; DB 6; Length 4721;
Best Local Similarity 97.2%; Pred. No. 9.8e-20;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TTGCCCACTCCCTCTCTGCGGCGCTCGCTCGCTCGGTGGGGCTGGGACCAAAAGGTCCGC 60
Db 1 TTGGCCACTCCCTCTCTATGCGCGCTCGCTCGCTCGGTGGGGCTGGGACCAAAAGGTCCGC 60

QY 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGAGCGAGCGAGCGAGCGAGGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGAGCGAGCGAGCGAGGTG 120

QY 121 GGCAACTCCATCACTAGGGGTA 142
Db 121 GGCAACTCCATCACTAGGGGTA 142

RESULT 9
CS073592 4721 bp DNA linear PAT 05-MAY-2005
LOCUS
DEFINITION
Sequence 180 from Patent WO200503321.
ACCESSION
CS073592
VERSION
CS073592.1 GI:63090489
KEYWORDS
unidentified
SOURCE
unidentified
ORGANISM
unclassified.

REFERENCE
1 Wilson,J.M., Gao,G., Alvira,M.R. and Vandenberghe,L.H.
Adeno-associated virus (aav) clades, sequences, vectors containing
same, and uses therefor
Patent: WO 200503321-A 180 14-APR-2005;
The Trustees of the University of Pennsylvania (US)

JOURNAL
Patent: WO 200503321-A 180 14-APR-2005;
The Trustees of the University of Pennsylvania (US)

FEATURES
source
Location/Qualifiers
1..4721
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/notes="adeno-associated virus serotype 7"

ORIGIN
Query Match 94.8%; Score 135.6; DB 6; Length 4721;
Best Local Similarity 97.2%; Pred. No. 9.8e-20;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TTGCCCACTCCCTCTCTGCGGCGCTCGCTCGCTCGGTGGGGCTGGGACCAAAAGGTCCGC 60
Db 1 TTGGCCACTCCCTCTCTATGCGCGCTCGCTCGCTCGGTGGGGCTGGGACCAAAAGGTCCGC 60

QY 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGAGCGAGCGAGCGAGGTG 120
Db 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGAGCGAGCGAGCGAGGTG 120

QY 121 GGCAACTCCATCACTAGGGGTA 142
Db 121 GGCAACTCCATCACTAGGGGTA 142

RESULT 10
AX753246 4721 bp DNA linear PAT 23-JUN-2003
LOCUS
DEFINITION
Sequence 1 from Patent EP1310571.
ACCESSION
AX753246
VERSION
AX753246.1 GI:32166105

ORIGIN
Query Match 95.9%; Score 137.2; DB 6; Length 4683;
Best Local Similarity 97.9%; Pred. No. 4.4e-20;
Matches 139; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTGCCCACTCCCTCTCTGCGGCGCTCGCTCGCTCGGTGGGGCTGGGACCAAAAGGTCCGC 60
Db 4683 TTGGCCACTCCCTCTCTATGCGCGCTCGCTCGCTCGGTGGGGCTGGGACCAAAAGGTCCGC 4624

QY 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGAGCGAGCGAGCGAGGTG 120
Db 4623 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGAGCGAGCGAGGTG 4564

QY 121 GGCAACTCCATCACTAGGGGTA 142
Db 4563 GGCAACTCCATCACTAGGGGTA 4542

RESULT 7
AR562507/c
LOCUS
DEFINITION
Sequence 19 from patent US 6759237.
ACCESSION
AR562507
VERSION
AR562507.1 GI:53976573
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unknown.

REFERENCE
1 (bases 1 to 4683)
Wilson,J.M. and Xiao,W.
Adeno-associated virus serotype 1 nucleic acid sequences, vectors
and host cells containing same
Patent: US 6759237-A 19 06-JUL-2004;
The Trustees of the University of Pennsylvania; Philadelphia, PA

JOURNAL
Patent: US 6759237-A 19 06-JUL-2004;
The Trustees of the University of Pennsylvania; Philadelphia, PA

FEATURES
source
Location/Qualifiers
1..4683
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 95.9%; Score 137.2; DB 6; Length 4683;
Best Local Similarity 97.9%; Pred. No. 4.4e-20;
Matches 139; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTGCCCACTCCCTCTCTGCGGCGCTCGCTCGCTCGGTGGGGCTGGGACCAAAAGGTCCGC 60
Db 4683 TTGGCCACTCCCTCTCTATGCGCGCTCGCTCGCTCGGTGGGGCTGGGACCAAAAGGTCCGC 4624

QY 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGAGCGAGCGAGCGAGGTG 120
Db 4623 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGAGCGAGCGAGGTG 4564

QY 121 GGCAACTCCATCACTAGGGGTA 142
Db 4563 GGCAACTCCATCACTAGGGGTA 4542

RESULT 8
CQ972062
LOCUS
DEFINITION
Sequence 7 from Patent WO2004108922.
ACCESSION
CQ972062
VERSION
CQ972062.1 GI:57163375
KEYWORDS
Adeno-associated virus
SOURCE
Adeno-associated virus
ORGANISM
Adeno-associated virus

REFERENCE
1 Rader,D.J. and Wilson,J.M.
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
Methods and compositions for lowering total cholesterol levels and
treatment of heart disease
```



```
KEYWORDS
SOURCE      Adeno-associated virus 7
ORGANISM    Adeno-associated virus 7
REFERENCE   1
AUTHORS     Gao,G., Wilson,J.M. and Alvira,M.
TITLE       A method of detecting and/or identifying adeno-associated virus
            (AVV) sequences and isolating novel sequences identified thereby
JOURNAL     Patent: EP 1310571-A 1 14-MAY-2003;
            The Trustees of The University of Pennsylvania (US)
FEATURES
  source
    1..4721
    /organism="Adeno-associated virus 7"
    /mol_type="unassigned DNA"
    /db_xref="taxon:202812"
ORIGIN
Query Match      94.8%; Score 135.6; DB 6; Length 4721;
Best Local Similarity 97.2%; Pred. No. 9.8e-20;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 TTGCCACCTCCCTCTCTCGCGCGCTCGCTCGCTGGGGCCCTGCGGACCAAAGGTCGCG 60
Db 1 TTGCCACCTCCCTCTCTATGCGCGCTCGCTCGCTGGGGCCCTGCGGACCAAAGGTCGCG 60
QY 61 AGACGGCAGAGTCTCTCTGCGCGGCCCAACCGAGCGAGCGCGCAGAGGAGGTG 120
Db 61 AGACGGCAGAGTCTCTCTGCGCGGCCCAACCGAGCGAGCGCGCAGAGGAGGTG 120
QY 121 GGCAACTCCATCACTAGGGGTA 142
Db 121 GGCAACTCCATCACTAGGGGTA 142
RESULT 11
AF513851
LOCUS       AF513851
DEFINITION  Adeno-associated virus 7 nonstructural protein and capsid protein
            genes, complete cds.
ACCESSION  AF513851
VERSION    AF513851.1 GI:22652859
KEYWORDS
SOURCE     Adeno-associated virus 7
ORGANISM   Adeno-associated virus 7
REFERENCE  1 (bases 1 to 4721)
AUTHORS    Gao,G.P., Alvira,M.R., Wang,L., Calcedo,R., Johnston,J. and
            Wilson,J.M.
TITLE      Novel adeno-associated viruses from rhesus monkeys as vectors for
            human gene therapy
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 99 (18), 11854-11859 (2002)
PUBMED    12192090
REFERENCE  2 (bases 1 to 4721)
AUTHORS    Alvira,M.R.
TITLE      Direct Submission
JOURNAL    Submitted (20-MAY-2002) Institute for Human Gene Therapy,
            University of Pennsylvania, M6.40 Maloney Bldg, 36th & Spruce Sts,
            Philadelphia, PA 19104, USA
FEATURES
  source
    1..4721
    /organism="Adeno-associated virus 7"
    /proviral
    /mol_type="genomic DNA"
    /specific_host="rhesus monkey"
    /db_xref="taxon:202812"
    334..2205
    /notes="similar to AAV2 Rep 78 protein"
    /codon_start=1
    /product="nonstructural protein"
    /protein_id="AA03854.1"
    /db_xref="GI:22652860"
    /translations="MFGFYEIVIKVPSDLDEHLPGISDSFVNVVAKEWELPPDSMD
    LNLIEQAPLTVAEKLRDFLQWRRVSKAPEALFFVQFEKGESYFHLHLVETGVKS"
CDS
  1..4721
  /note="similar to AAV2 VP1 protein"
  /codon_start=1
  /product="capsid protein"
  /protein_id="AA03855.1"
  /db_xref="GI:22652861"
  /translation="MAADGYLPDWLEDNLSGIREWMDLKPQAPKPKANQQQDNGRGG
  LVLPGYKYLGPNGLDKGEVNAADAALEHDKAYDQQLKAGDNPYLRYNHADAEPQE
  RLOEDTSGGNLGRAVFOAKKEVLEPLGLIVERGAKTAPAKRPEVPSPPORSDDSTGI
  GKGGQAPARKLNFQGTGDSVDPDPLGEPAPAPSSVSGCTVAAGGAPWADNNEG
  ADGVNAGSNHCHDSITWGLDRVITSTRTWTALPTNNHLYKQISSETAGSTNDNTYFG
  YSTPWGYDFDNRFCHFSRDMQRLNNNNWGRPKKLPKLFNIQVKEVTINDGVTTI
  ANNLTSTIQVFSDSYQLPYVLGSAHQGLPPFPADVFMIPOYGLTLNNGSQSVGRS
  SPYCLEYFPQMLRGTNNNFESYSPEDVPFHSLSQHSQSLDLMNPLIDQYLYYLART
  OSNPGGTAGNRELQYQGGPSTMAQKWLPGCFRQORVSKTLDONNNSNFAWTGA
  TKYHLGNLSLNVPGVAMATKDDDRPPSSGVLIFGKTGATNKTLENLVMTNEE
  IRTPNPVATEEYGIYSSNLQAANTAAQTQVNNQAGALPGVWQNRDVIQGGPIWAKLP
  HTDGNFHPSPLMGGFGLKHPQIILIKNTVPANPEVFTPAKFASFITQYSTGQVSV
  ETEWELQKENSKRWNPEIQYTSNFEKQTGVDPAVDSQGVYSEPRPIGTRYLRNL"
```

```

MVLGRFLSQIREKLVQTIYRGVEPTLPNFAVTKTRNGAGGKNKVVDCEYIPNYLLPK
TQPELQWATNMEEYISACLNAERKRLVQAQHLTVSQTQEQNKENLNPNSDAPVIRS
KTSARYMELVGNLVDRGITSEKQMTIQEODASYISFNAASNSRSQIKAAALDNAGKIAL
TKSADPYLVGPSLPADIKTNRIYRIELNGDYPAYAGSVFLGWAQKTFKRNITWLFQ
PATTKTNIAEIAHAVPFYGCNVNNTNFPNDCVDMVIMWEGKMTAKVVESSAKA
ILGGSKRVQDKQSSAOIDPTPVITNTNMCAVIDGNSSTTFHQQDQRMFKFEL
TRLEHDFGKVTQKQVFFRPAWSHVTVEAHFVYVRKGASKRPAQDDADISEPKRA
CPSVADPSTDAEGAPVDFADRYQNKSRHAGMIQMLFPCKTCERMNQNFNICFTHGV
RDCLECFPGVSESPVVRKTKYRKLCAIHLGLRAPEIACASACDLVNVLDLDCVSEQ"
2222..4435
/note="similar to AAV2 VP1 protein"
/codon_start=1
/product="capsid protein"
/protein_id="AA03855.1"
/db_xref="GI:22652861"
/translation="MAADGYLPDWLEDNLSGIREWMDLKPQAPKPKANQQQDNGRGG
LVLPGYKYLGPNGLDKGEVNAADAALEHDKAYDQQLKAGDNPYLRYNHADAEPQE
RLOEDTSGGNLGRAVFOAKKEVLEPLGLIVERGAKTAPAKRPEVPSPPORSDDSTGI
GKGGQAPARKLNFQGTGDSVDPDPLGEPAPAPSSVSGCTVAAGGAPWADNNEG
ADGVNAGSNHCHDSITWGLDRVITSTRTWTALPTNNHLYKQISSETAGSTNDNTYFG
YSTPWGYDFDNRFCHFSRDMQRLNNNNWGRPKKLPKLFNIQVKEVTINDGVTTI
ANNLTSTIQVFSDSYQLPYVLGSAHQGLPPFPADVFMIPOYGLTLNNGSQSVGRS
SPYCLEYFPQMLRGTNNNFESYSPEDVPFHSLSQHSQSLDLMNPLIDQYLYYLART
OSNPGGTAGNRELQYQGGPSTMAQKWLPGCFRQORVSKTLDONNNSNFAWTGA
TKYHLGNLSLNVPGVAMATKDDDRPPSSGVLIFGKTGATNKTLENLVMTNEE
IRTPNPVATEEYGIYSSNLQAANTAAQTQVNNQAGALPGVWQNRDVIQGGPIWAKLP
HTDGNFHPSPLMGGFGLKHPQIILIKNTVPANPEVFTPAKFASFITQYSTGQVSV
ETEWELQKENSKRWNPEIQYTSNFEKQTGVDPAVDSQGVYSEPRPIGTRYLRNL"
ORIGIN
Query Match      94.8%; Score 135.6; DB 13; Length 4721;
Best Local Similarity 97.2%; Pred. No. 9.8e-20;
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 TTGCCACCTCCCTCTCTCGCGCGCTCGCTCGCTGGGGCCCTGCGGACCAAAGGTCGCG 60
Db 1 TTGCCACCTCCCTCTATGCGCGCTCGCTCGCTGGGGCCCTGCGGACCAAAGGTCGCG 60
QY 61 AGACGGCAGAGTCTCTGCTGCGCGGCCCAACCGAGCGAGCGCGCAGAGGAGGTG 120
Db 61 AGACGGCAGAGTCTCTGCTGCGCGGCCCAACCGAGCGAGCGCGCAGAGGAGGTG 120
QY 121 GGCAACTCCATCACTAGGGGTA 142
Db 121 GGCAACTCCATCACTAGGGGTA 142
RESULT 12
BD242766/c
LOCUS       BD242766
DEFINITION  Adeno-associated virus serum type 1 nucleic acid sequence, vector
            and host cell containing the same.
ACCESSION  BD242766
VERSION    BD242766.1 GI:33052536
KEYWORDS   JP 2002529098-A/1.
SOURCE     unidentified
ORGANISM   unidentified
REFERENCE  1 (bases 1 to 4718)
AUTHORS    Wilson,J.M. and Xiao,W.
TITLE      Adeno-associated virus serum type 1 nucleic acid sequence, vector
            and host cell containing the same
JOURNAL    Patent: JP 2002529098-A 1 10-SEP-2002;
            THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
COMMENT    OS AAV-1
            PN JP 2002529098-A/1
            PD 10-SEP-2002
            PF 02-NOV-1999 JP 2000581227
            PR 05-NOV-1998 US 60/107114
            PI JAMES M WILSON, WEIDONG XIAO
            PC C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC
            C12N1/21,
            PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC
```

Adeno-associated virus serum type 1 nucleic acid sequence, CC  
vector and host  
CC cell containing the same  
FH Key Location/Qualifiers  
FT CDS (335)..(2206)  
FT CDS (2223)..(4430).  
Location/Qualifiers  
1..4718  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 83.2%; Score 119; DB 6; Length 4718;  
Best Local Similarity 89.5%; Pred. No. 3.7e-16;  
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 1 TTGCCACTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCTCGCGACCAAGGTCCGC 60  
Db 4718 TTGCCACTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCTCGCGACCAAGGTCTGC 4659  
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 120  
Db 4658 CGTCTCGGACCTTTGGTCCGAGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 4599  
QY 121 GGCACTCCATCACTAGGGGTAA 143  
Db 4598 GGCACTCCATCACTAGGGGTAA 4576  
RESULT 13  
AR562498/C  
LOCUS 4718 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 1 from patent US 6759237.  
ACCESSION AR562498  
VERSION AR562498.1 GI:53976564  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4718)  
AUTHORS Wilson,J.M. and Xiao,W.  
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors and host cells containing same  
JOURNAL Patent: US 6759237-A 1 08-JUL-2004;  
The Trustees of the University of Pennsylvania; Philadelphia, PA  
FEATURES  
source 1..4718  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 83.2%; Score 119; DB 6; Length 4718;  
Best Local Similarity 89.5%; Pred. No. 3.7e-16;  
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 1 TTGCCACTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCTCGCGACCAAGGTCCGC 60  
Db 4718 TTGCCACTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCTCGCGACCAAGGTCTGC 4659  
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 120  
Db 4658 CGTCTCGGACCTTTGGTCCGAGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 4599  
QY 121 GGCACTCCATCACTAGGGGTAA 143  
Db 4598 GGCACTCCATCACTAGGGGTAA 4576  
RESULT 14  
AX753251/C  
LOCUS 4718 bp DNA linear PAT 23-JUN-2003  
DEFINITION Sequence 6 from Patent EP1310571.

AX753251 AX753251.1 GI:32166108  
KEYWORDS  
SOURCE Adeno-associated virus 1  
ORGANISM Adeno-associated virus 1  
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  
REFERENCE 1  
AUTHORS Gao,G., Wilson,J.M. and Alvira,M.  
TITLE A method of detecting and/or identifying adeno-associated virus (AVV) sequences and isolating novel sequences identified thereby  
JOURNAL Patent: EP 1310571-A 6 14-MAY-2003;  
The Trustees of the University of Pennsylvania (US)  
FEATURES  
Location/Qualifiers  
source 1..4718  
/organism="Adeno-associated virus 1"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:85106"  
ORIGIN  
Query Match 83.2%; Score 119; DB 6; Length 4718;  
Best Local Similarity 89.5%; Pred. No. 3.7e-16;  
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 1 TTGCCACTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCTCGCGACCAAGGTCCGC 60  
Db 4718 TTGCCACTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCTCGCGACCAAGGTCTGC 4659  
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 120  
Db 4658 CGTCTCGGACCTTTGGTCCGAGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 4599  
QY 121 GGCACTCCATCACTAGGGGTAA 143  
Db 4598 GGCACTCCATCACTAGGGGTAA 4576  
RESULT 15  
AF063497/C  
LOCUS 4718 bp DNA linear VRL 27-APR-1999  
DEFINITION Adeno-associated virus 1, complete genome.  
ACCESSION AF063497  
VERSION AF063497.1 GI:4689096  
KEYWORDS  
SOURCE Adeno-associated virus 1  
ORGANISM Adeno-associated virus 1  
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  
REFERENCE 1 (bases 1 to 4718)  
AUTHORS Xiao,W., Chirmule,N., Berta,S.C., McCullough,B., Gao,G. and Wilson,J.M.  
TITLE Gene therapy vectors based on adeno-associated virus type 1  
JOURNAL J. Virol. 73 (5), 3994-4003 (1999)  
PUBMED 10196295  
REFERENCE 2 (bases 1 to 4718)  
AUTHORS Xiao,W. and Wilson,J.M.  
TITLE Direct Submission  
JOURNAL Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601 Spruce Street, Philadelphia, PA 19104, USA  
FEATURES  
Location/Qualifiers  
source 1..4718  
/organism="Adeno-associated virus 1"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:85106"  
/note="AAV1"  
335..2206  
/codon\_start=1  
/product="nonstructural protein"  
/protein\_id="AAD27758.1"  
/db\_xref="GI:4689098"  
/translation="MPGYEIVIKVPSDLDEHLPGISDFVSWAEKEWELPPDSMD  
LNLIEQALPTVAELKLDLFVOMRRVSKAPEALFVQFEKGESYFHLHLLVSTTVKRS  
MYLGRFLSQRIDKLQVTYRGITPLPNFVAVTKTRNGAGGKGVDBCYIYNYLLPK  
TOPELQAWNTMNEEYISACLNLAEKRLVAOHLTHVSQTQEQNKENLNPNSDAPVRS  
KTSARYMELVGLVDRGITSEKQWIQEDQASVYSFNAASNSRSQIKALDNLNAGKIMAL  
CDS

TKSADYLVGPAPPADIKTNRIYRIELENGYEPAYAGSVFLGWAQKRFGRNTIWLFG  
PATTKTNIAEIAHAHAFYGCWNNTNENFPENDCVDKMVIWEEGKMTAKRVESAKA  
ILGSKVRVDQKCKSAQIDPTPIVTISNTNCVIDGNTSTTFHQQLQDRMFKEFEL  
TRRLBHDGKVTQKQVKEPFRWAQDHTEVAHEFYVRKGANKRPAPDDADKSEPKRA  
CPSVADPSTSDAEGAPVDPADRYQNKCSFHAGMLOMLPCKTCERMNONFNICFTHGT  
RDCSECFQGVSESQPVVRKRYRKUICAIHLLGRAPETACSACDLVNVDDLDDCVSEQ"  
2223. .4433  
/codon\_start=1  
/product="capsid protein"  
/protein\_id="AAD27757.1"  
/db\_xref="GI:4689097"  
/translation="MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQKQDDGRG  
LVLPGYKILGPPNGLDKGEFVNAADAALAHDKAYDQQLKAGDNPILRYNHADAEFQE  
RLQEDTSFGNIGRAVFOAKRVLPLGLVEEGAKTAPGKRKPVQSPQEPDSSGIG  
KTQOPAKRRLNFGOTGSESVDPQLCEPPATPAAGPTTMASSGGAPMADNNEGA  
DGVNAGSNHCHDSIWLGDRTVTSTRTWALPTNNHLYKQISSASTGASNDNHVFGY  
STPMGYFDNRPHCHPSPRDWORLLNNWGFHFKRLNFKLENIQVKEVTNDGVTTIA  
NNLTSTVQVFSSEYQLPYVLGSAHQCLPPFPADVPMIPQYGYLTLANGSAQGRSS  
FYCLEYFQSQMLRTGNFTSYTFEEVPHSHSQSLDRMLNPLIDQYLYLNRTO  
NQSGAQNKDLLFSRGSFAGMSVQPKNMLPGPCYRQQRVSKTKTDNNNSNFTWTGASK  
YNLNGRESTINPCTAMASHKDDKFFPMGSGVMIIFGESAGASNTALDNVMTDBEEI  
KATNPVATERFCTVAVNPOSSSTDPTADGVHAMGALPGMWODRDVYLQGPWAKIPH  
TDGHFHPSLMGFGGLKNPPQIILKNTPVPANPPAEFSATKFASTIQYSTGQVSVE  
IEWELQKENSKNWNPVQYTSNYSKANSVDFTVDNNGLYTEPRPIGTRYLTRPL"

ORIGIN

Query Match 83.2%; Score 119; DB 13; Length 4718;  
Best Local Similarity 89.5%; Pred. No. 3.7e-16;  
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
  
QY 1 TTGCCCCACTCCTCTCTGCGCGCTCGCTCGCTGGTGGGGCCTGCGGACCAAGGTCCGC 60  
Db 4718 TTGCCCCACTCCTCTCTGCGCGCTCGCTCGCTGGTGGGGCCTGCGGACCTCTGC 4659  
  
QY 61 AGACGGCAGAGCTCTCTCTGCGGCGCCACCGAGCGAGCGCGGAGAGGAGTG 120  
Db 4658 CGTCTCGGACCTTTTGGTCCGAGCGCCCAACCGAGCGAGCGCGGAGGAGTG 4599  
  
QY 121 GGCACTCCATCACTAGGGGTAA 143  
Db 4598 GGCACTCCATCACTAGGGGTAA 4576

Search completed: November 29, 2005, 00:05:25  
Job time : 524.566 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 08:46:14 ; Search time 66.9124 Seconds  
(without alignments)  
14243.261 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_1\_143

Perfect score: 143

Sequence: 1 ttgccactccctctctgcg.....aactccatcactagggttaa 143

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*
- 14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	100.0	143	3	AAD00781
2	143	100.0	4718	3	AAD00772
3	143	100.0	4718	10	AD876507
4	143	100.0	4718	10	ADL13984
5	143	100.0	4718	12	ADG39758
6	138.8	97.1	4683	4	AAF23749
7	138.8	97.1	4683	10	ADL13983
8	138.8	97.1	4683	12	ADG39763
9	135.6	94.8	4721	10	AD876502
10	135.6	94.8	4721	12	ADG39764
11	135.6	94.8	4721	14	ADV67509
12	135.6	94.8	4721	14	ADZ27030
13	119	83.2	4718	3	AAD00772
14	119	83.2	4718	10	AD876507
15	119	83.2	4718	10	ADL13984
16	119	83.2	4718	12	ADG39758
17	113.2	79.2	4722	4	AAF23748
18	113.2	79.2	4722	12	ADG39760
19	113.2	79.2	4726	10	AD876509

20	113.2	79.2	4726	10	ADL13985
21	113.2	79.2	4726	12	ADG39759
22	112.2	78.5	4722	4	AAF23748
23	112.2	78.5	4722	12	ADG39760
24	111.6	78.0	4721	10	AD876502
25	111.6	78.0	4721	12	ADG39764
26	111.6	78.0	4721	14	ADV67509
27	111.6	78.0	4721	14	ADZ27030
28	107.4	75.1	145	2	AAQ41448
29	107.4	75.1	145	2	AAT03385
30	107.4	75.1	145	6	ABX34295
31	107.4	75.1	145	6	ABSG6984
32	107.4	75.1	145	8	ACC58491
33	107.4	75.1	145	9	ACF35876
34	107.4	75.1	145	13	ADM93250
35	107.4	75.1	146	8	ABX93568
36	107.4	75.1	165	2	AAQ66769
37	107.4	75.1	165	2	AAT49462
38	107.4	75.1	165	4	AAD03535
39	107.4	75.1	165	6	ABSG6986
40	107.4	75.1	165	8	ABV77279
41	107.4	75.1	170	12	ADL91189
42	107.4	75.1	207	6	AA44621
43	107.4	75.1	272	3	AZ47166
44	107.4	75.1	272	4	AAD03534
45	107.4	75.1	955	6	AAD37254

## ALIGNMENTS

RESULT 1

AAD00781

ID AAD00781 standard; DNA; 143 BP.

XX AAD00781;

XX 08-SEP-2000 (first entry)

XX Adeno-associated virus serotype 1 5' inverted terminal repeat.

XX Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;

KW cap protein; recombinant viral vector; Gene delivery; gene therapy;

KW vaccine; transgene; inverted terminal repeat; ITR; ss.

XX Adeno associated virus serotype 1.

XX Key Location/Qualifiers

FT misc\_binding 1..41

FT /tag= a

FT /bound\_moiety= "binds to nucleotides 85..125"

FT stem\_loop 42..62

FT /tag= b

FT stem\_loop 64..84

FT /tag= c

FT misc\_binding 85..125

FT /tag= d

FT /bound\_moiety= "binds to nucleotides 41..1"

XX WC200028061-A2.

XX 18-MAY-2000.

XX 02-NOV-1999; 99NO-US025694.

XX 05-NOV-1998; 98US-0107114P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Wilson JM, Xiao W;

XX WPI; 2000-376571/32.

XX

PT	Novel adeno-associated virus serotype 1 polynucleotide useful for preparation of medicament for delivery of a transgene to a host.
xx	
P8	Example 2; Fig 2; 108pp; English.
xx	
CC	The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA which is characterised by two inverted terminal repeats (ITR) and open reading frames for rep and capsid (cap) proteins. The rep reading frame encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1 sequence or its fragments particularly ITRs, rep and cap coding regions, are useful in production of recombinant viral vectors for gene delivery. These vectors can be used as gene therapy vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does not induce the formation of neutralising antibodies specific to any serotype of AAV hence is useful for transforming host cells, and in the preparation of a medicament for the delivery of transgene to a host. The present sequence is a 5' ITR of AAV-1 DNA which is useful in the production of recombinant viral vector. The ITR forms a T-shaped hairpin structure
xx	
QQ	Sequence 143 BP; 24 A; 50 C; 48 G; 21 T; 0 U; 0 Other;
	Query Match 100.0%; Score 143; DB 3; Length 143;
	Best Local Similarity 100.0%; Pred. No. 2.3e-28;
	Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 TTGCCACATCCCTCTCTGCGGCTCGCTCGGTCGGTGGGGCTCGCGACCAAAAGGTCGCG 60
DB	1 TTGCCACATCCCTCTCTGCGGCTCGCTCGGTCGGTGGGGCTCGCGACCAAAAGGTCGCG 60
QY	61 AGAGCGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGGGAGGTG 120
DB	61 AGAGCGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGGGAGGTG 120
QY	121 GGCAACTCCATCACTAGGGTTAA 143
DB	121 GGCAACTCCATCACTAGGGTTAA 143
xx	
RESULT 2	
AAD00772	
ID	AAD00772 standard; DNA; 4718 BP.
XX	
AC	AAD00772;
XX	
DT	08-SEP-2000 (first entry)
XX	
DE	Adeno-associated virus serotype 1 DNA.
XX	
KW	Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
KW	cap protein; recombinant viral vector; gene delivery; gene therapy;
KW	vaccine; transgene; ss.
XX	
OS	Adeno associated virus serotype 1.
XX	
PH	Key Location/Qualifiers
FT	repeat_unit 1..143
FT	/*tag= a
FT	/label= 5' ITR
FT	/note= "Inverted terminal repeat which is capable of forming T-shaped hairpin structure"
FT	protein_bind 89..110
FT	/*tag= b
FT	/bound_moiety= "Rep protein"
FT	misc_feature 124..125
FT	/*tag= c
FT	/note= "Terminal resolute site (TRS)"
FT	misc_binding 219..226
FT	/*tag= d
FT	/bound_moiety= "USF"
FT	/note= "E box"
FT	promoter 236..299
FT	/*tag= e

PF 02-NOV-1999; 99WO-US025694.  
XX  
PR 05-NOV-1998; 98US-0107114P.  
PR  
PA (UYPE-) UNIV PENNSYLVANIA.  
XX  
XX  
PI Wilson JM, Xiao W;  
XX  
DR WPI: 2000-376571/32.  
DR P-PSDB; AAY71161, AAY71164, AAY71165, AAY71166, AAY71167, AAY71168,  
DR AAY71169.  
XX  
XX Novel adeno-associated virus serotype 1 polynucleotide useful for  
PT preparation of medicament for delivery of a transgene to a host.  
XX  
XX Claim 1; Fig 1; 108pp; English.  
XX  
XX The present sequence is an adeno-associated virus serotype 1 (AAV-1) DNA  
CC characterised by two inverted terminal repeats (ITR) and open reading  
CC frames for rep and capsid (cap) proteins. The rep reading frame encodes  
CC four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading  
CC frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1  
CC sequence or its fragments particularly ITRs, rep and cap coding regions,  
CC are useful in production of recombinant viral vectors for gene delivery.  
CC These vectors can be used as gene therapy vectors, vaccine vectors or  
CC antineoplastic delivery vectors. The AAV-1 does not induce the formation of  
CC neutralising antibodies specific to any serotype of AAV hence is useful  
CC for transforming host cells, and in the preparation of a medicament for  
CC the delivery of transgene to a host  
XX  
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;  
Query Match 100.0%; Score 143; DB 3; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 2.8e-28;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCCCTCGCGACCAAGGTCCGC 60  
Db 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCCCTCGCGACCAAGGTCCGC 60  
QY 61 AGACGGCAGAGCTCTGCTCTGCGGGCCCCACCGAGCGAGCGCGCGCAGAGGGAGTG 120  
Db 61 AGACGGCAGAGCTCTGCTCTGCGGGCCCCACCGAGCGAGCGCGCGCAGAGGGAGTG 120  
QY 121 GGCAACTCCATCACTAGGGGTAA 143  
Db 121 GGCAACTCCATCACTAGGGGTAA 143  
RESULT 3  
ADE76507  
ID ADE76507 standard; DNA; 4718 BP.  
XX  
AC ADE76507;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Adeno-associated virus (AAV) related DNA, SEQ ID No 6.  
XX  
XX adeno-associated virus; AAV; cytostatic; antipsoriatic; antirheumatic;  
KW antiarthritic; neuroprotective; antidiabetic; antithyroid;  
KW dermatological; antiinflammatory; gene therapy; vaccine;  
KW hyperproliferative; cancer; psoriasis; autoimmune disease;  
KW rheumatoid arthritis; multiple sclerosis; diabetes;  
KW autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.  
XX  
OS Adeno-associated virus 1.  
XX  
PN EF1310571-A2.  
XX  
XX 14-MAY-2003.  
XX  
PF 12-NOV-2002; 2002EP-00257826.

XX 13-NOV-2001; 2001US-0350607P.  
PR 17-DEC-2001; 2001US-0341117P.  
PR 01-MAY-2002; 2002US-0377066P.  
PR 05-JUN-2002; 2002US-0386675P.  
XX  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX  
XX Gao G, Wilson JM, Alvira M;  
XX WPI: 2003-450984/43.  
DR  
XX  
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.  
PT preventing or treating hyperproliferative or autoimmune diseases,  
PT comprises subjecting a sample having a DNA to amplification via  
PT polymerase chain reaction.  
XX  
XX Claim 14; SEQ ID NO 6; 419pp; English.  
XX  
XX The invention relates to a novel method for detecting adeno-associated  
CC virus (AAV) sequences in a sample, which comprises subjecting a sample  
CC containing a DNA to amplification via a polymerase chain reaction (PCR).  
CC The AAV sequence have the following activities: cytostatic,  
CC antipsoriatic, antirheumatic, antarthritic, neuroprotective,  
CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV  
CC sequence can be used in gene therapy or as part of a vaccine to treat  
CC disorders. The method is useful in detecting and/or identifying AAV  
CC sequences and isolating novel sequences that are identified. The  
CC sequences may be used e.g. for preventing or treating hyperproliferative  
CC conditions such as cancers and psoriasis, and other autoimmune diseases  
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune  
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence  
CC represents an AAV related DNA sequence of the invention.  
XX  
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;  
Query Match 100.0%; Score 143; DB 10; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 2.8e-28;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCCCTCGCGACCAAGGTCCGC 60  
Db 1 TTGCCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCCCTCGCGACCAAGGTCCGC 60  
QY 61 AGACGGCAGAGCTCTGCTCTGCGGGCCCCACCGAGCGAGCGCGCGCAGAGGGAGTG 120  
Db 61 AGACGGCAGAGCTCTGCTCTGCGGGCCCCACCGAGCGAGCGCGCGCAGAGGGAGTG 120  
QY 121 GGCAACTCCATCACTAGGGGTAA 143  
Db 121 GGCAACTCCATCACTAGGGGTAA 143  
RESULT 4  
ADL13984  
ID ADL13984 standard; DNA; 4718 BP.  
XX  
XX ADL13984;  
XX  
XX 06-MAY-2004 (first entry)  
XX  
XX Adeno-associated virus serotype 1 complete DNA.  
XX  
XX ss; cytostatic; neuroprotective; antiinflammatory; gene therapy;  
KW expression construct; adeno-associated virus;  
KW integration efficiency element; inverted terminal repeat; integration;  
KW chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;  
KW retinoblastoma; inflammatory disease; arthritis;  
KW neurodegenerative disease.  
XX  
XX Adeno-associated virus 1.  
XX  
XX WO2003087334-A2.  
PN



XX PD 23-OCT-2003.  
 XX PF 09-APR-2003; 2003WO-US011191.  
 XX PR 09-APR-2002; 2002US-0371044P.  
 XX PA (CORR ) CORNELL RES FOUND INC.  
 XX PI Falck-Pedersen ES, Philipott N;  
 XX DR WPI; 2003-833723/77.  
 XX PT New expression construct comprising a nucleic acid sequence encoding an  
 PT adeno-associated virus integration efficiency element, useful for  
 PT treating cancer e.g. lung cancer or colon cancer or inflammatory disease  
 PT e.g. arthritis.  
 XX PT Disclosure; SEQ ID NO 3; 62pp; English.  
 XX PS The invention relates to an expression construct comprising a nucleic  
 XX acid sequence encoding an adeno-associated virus integration efficiency  
 CC element (AAV IRE), which is devoid of AAV inverted terminal repeats (AAV  
 CC ITRs) and site-specifically integrates into a host cell chromosome when  
 CC provided to the host cell in conjunction with an AAV Rep protein. The  
 CC expression construct can be used as a therapeutic factor for treating a  
 CC mammal for a pathologic state which is cancer, including lung cancer,  
 CC colon cancer, renal cancer, anal cancer, bile duct cancer, bladder  
 CC cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,  
 CC cervical cancer, lymphoma, endometrial cancer, esophageal cancer,  
 CC gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,  
 CC liver cancer, multiple myeloma, neuroblastoma, ovarian cancer, pancreatic  
 CC cancer, prostatic cancer, retinoblastoma, skin cancer, stomach cancer,  
 CC testicular cancer, thymus cancer or thyroid cancer. Other pathologic  
 CC state includes inflammatory disease (arthritis), neurodegenerative  
 CC disease, a disease of an organ attributed to the presence of increased or  
 CC decreased level of a particular gene product(s). This sequence  
 CC corresponds to the AAV serotype 1 complete DNA including the IRE  
 CC sequence.  
 XX PT Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 143; DB 10; Length 4718;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-28;  
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCTGCGGACCAAGGTCCGC 60  
 Db 1 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCTGCGGACCAAGGTCCGC 60  
 QY 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGCGGAGGAGTG 120  
 Db 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGCGGAGGAGTG 120  
 QY 121 GGCAACTCCATCACTAGGGGTAA 143  
 Db 121 GGCAACTCCATCACTAGGGGTAA 143  
 RESULT 5  
 ADG39758  
 ID ADG39758 standard; DNA; 4718 BP.  
 XX AC ADG39758;  
 XX DT 11-MAR-2004 (first entry)  
 XX DE AAV-1 genomic DNA sequence SEQ ID NO:20.  
 XX parvovirus; rep; cap; DNA binding domain; capsid interacting domain;  
 KW recombinant hybrid parvovirus particle;  
 KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;  
 KW gene; ds.

XX OS Adeno-associated virus 1.  
 XX PN WO2003104392-A2.  
 XX PD 18-DEC-2003.  
 XX PF 02-DEC-2002; 2002WO-US038423.  
 XX PR 18-DEC-2001; 2001US-0341919P.  
 XX PA (UYNC-) UNIV NORTH CAROLINA.  
 XX PI Samulski RJ, Rabinowitz JE;  
 XX DR WPI; 2004-062324/06.  
 XX PT New polynucleotides comprising parvovirus rep coding sequences and  
 PT parvovirus cap coding sequences, useful in producing higher stocks of  
 PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a  
 PT subject.  
 XX PT Disclosure; SEQ ID NO 20; 115pp; English.  
 XX PS The present invention describes a polynucleotide (I), comprising  
 CC parvovirus rep coding sequences and parvovirus cap coding sequences. The  
 CC rep coding sequences encodes a DNA binding domain from a first  
 CC parvovirus, and a capsid interacting domain from a parvovirus different  
 CC from the first parvovirus. The cap coding sequence comprises sequences  
 CC from the different parvovirus. Also described: (1) a vector comprising  
 CC (I); (2) a cell comprising (I), or parvovirus rep coding sequences and  
 CC parvovirus cap coding sequences, where the rep coding sequences encode a  
 CC DNA binding domain from a first parvovirus and a capsid interacting  
 CC domain from a parvovirus different from the first parvovirus, the cap  
 CC coding sequences comprises sequences from the different parvovirus, and  
 CC the rep coding sequences are stably integrated into the genome of the  
 CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno  
 CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in  
 CC gene therapy. The polynucleotide (I) can be used in producing higher  
 CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in  
 CC the delivery of nucleic acids having biological effect to treat or  
 CC ameliorate the symptoms associated with any disorder related to gene  
 CC expression. The polynucleotide may be used to produce a parvovirus vector  
 CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.  
 CC The parvovirus vector may also be used to provide an antisense nucleic  
 CC acid to a cell in vitro or in vivo, or in diagnostic and screening  
 CC methods. The present sequence is used in the exemplification of the  
 CC present invention.  
 XX PT Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 143; DB 12; Length 4718;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-28;  
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCTGCGGACCAAGGTCCGC 60  
 Db 1 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCTGCGGACCAAGGTCCGC 60  
 QY 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGCGGAGGAGTG 120  
 Db 61 AGACGGCAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGCGGAGGAGTG 120  
 QY 121 GGCAACTCCATCACTAGGGGTAA 143  
 Db 121 GGCAACTCCATCACTAGGGGTAA 143  
 RESULT 6  
 AAF23749/c  
 ID AAF23749 standard; DNA; 4683 BP.  
 XX AC AAF23749;  
 XX

XX 28-MAR-2001 (first entry)  
XX AAV6 DNA sequence.  
XX AAV6; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;  
KW atherosclerosis; sickle cell anemia; thalassemia;  
KW blood clotting disorder; diabetes; ss.  
XX Adeno associated virus.  
XX US6156303-A.  
XX 05-DEC-2000.  
XX 11-JUN-1997; 97US-00873168.  
XX 11-JUN-1997; 97US-00873168.  
XX (UNIW ) UNIV WASHINGTON.  
XX Russell DW, Rutledge EA;  
XX WPI; 2001-060164/07.  
XX Adeno-associated virus serotype 6 and viral vector derived from it for  
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency  
PT syndrome, sickle cell anemia, thalassemia and diabetes.  
XX Claim 1; Fig 1; 50pp; English.  
XX The present invention relates to adeno-associated virus serotypes. The  
XX present sequence is the DNA sequence of one such serotype (AAV6). AAV6  
XX can be used to construct AAV viral vectors for use in gene therapy for a  
XX range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,  
XX sickle cell anemia, thalassemia, blood clotting disorders and diabetes.  
XX The AAV viral vectors have increased transduction efficiency of a  
XX particular host cell as the AAV virion containing the AAV vector genome  
XX can be modified to express a capsid protein of an AAV serotype that  
XX transduces the selected host cell  
XX  
XX Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 U; 0 Other;  
Query Match 97.1%; Score 138.8; DB 4; Length 4683;  
Best Local Similarity 98.6%; Pred. No. 3.5e-27;  
Matches 140; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTGGTGGGCGCTCGGACCAAAAGGTCCGC 60  
Db 4683 TTGCCCACTCCCTCTATGCGGCTCGCTCGCTGGTGGGCGCTCGGACCAAAAGGTCCGC 4624  
QY 61 AGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 120  
Db 4623 AGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 4564  
QY 121 GGCAACTCCATCACTAGGGGTA 142  
Db 4563 GGCAACTCCATCACTAGGGGTA 4542  
RESULT 7  
ADL13983/c  
ID ADL13983 standard; DNA; 4683 BP.  
XX ADL13983;  
XX 06-MAY-2004 (first entry)  
XX Adeno-associated virus serotype 6 complete DNA.  
XX ss; cytostatic; neuroprotective; antiinflammatory; gene therapy;  
KW expression construct; adeno-associated virus;  
KW integration efficiency element; inverted terminal repeat; integration;

KW Chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;  
KW retinoblastoma; inflammatory disease; arthritis;  
KW neurodegenerative disease.  
XX Adeno-associated virus 6.  
XX WO2003087334-A2.  
XX 23-OCT-2003.  
XX 09-APR-2003; 2003WO-US011191.  
XX 09-APR-2002; 2002US-0371044P.  
XX (CORR ) CORNELL RES FOUND INC.  
XX Falck-Pedersen ES, Philpott N;  
XX WPI; 2003-833723/77.  
XX New expression construct comprising a nucleic acid sequence encoding an  
PT adeno-associated virus integration efficiency element, useful for  
PT treating cancer e.g. lung cancer or colon cancer or inflammatory disease  
PT e.g. arthritis.  
XX Disclosure; SEQ ID NO 2; 62pp; English.  
XX The invention relates to an expression construct comprising a nucleic  
XX acid sequence encoding an adeno-associated virus integration efficiency  
XX element (AAV IEE), which is devoid of AAV inverted terminal repeats (AAV  
XX ITRs) and site-specifically integrates into a host cell chromosome when  
XX provided to the host cell in conjunction with an AAV Rep protein. The  
XX expression construct can be used as a therapeutic factor for treating a  
XX mammal for a pathologic state which is cancer, including lung cancer,  
XX colon cancer, renal cancer, anal cancer, bile duct cancer, bladder  
XX cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,  
XX cervical cancer, lymphoma, endometrial cancer, esophageal cancer,  
XX gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,  
XX liver cancer, multiple myeloma, neuroblastoma, ovarian cancer, pancreatic  
XX cancer, prostatic cancer, retinoblastoma, skin cancer, stomach cancer,  
XX testicular cancer, thymus cancer or thyroid cancer. Other pathologic  
XX state includes inflammatory disease (arthritis), neurodegenerative  
XX disease, a disease of an organ attributed to the presence of increased or  
XX decreased level of a particular gene product(s). This sequence  
XX corresponds to the AAV serotype 6 complete DNA including the IEE  
XX sequence.  
XX Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 U; 0 Other;  
Query Match 97.1%; Score 138.8; DB 10; Length 4683;  
Best Local Similarity 98.6%; Pred. No. 3.5e-27;  
Matches 140; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTGGTGGGCGCTCGGACCAAAAGGTCCGC 60  
Db 4683 TTGCCCACTCCCTCTATGCGCGCTCGCTCGCTGGTGGGCGCTCGGACCAAAAGGTCCGC 4624  
QY 61 AGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 120  
Db 4623 AGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 4564  
QY 121 GGCAACTCCATCACTAGGGGTA 142  
Db 4563 GGCAACTCCATCACTAGGGGTA 4542  
RESULT 8  
ADG39763/c  
ID ADG39763 standard; DNA; 4683 BP.  
XX ADG39763;  
XX 11-MAR-2004 (first entry)

```
XX AAV-6 genomic DNA sequence SEQ ID NO:25.
DE parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
XX recombinant hybrid parvovirus particle;
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
KW gene; ds.
XX
OS Adeno-associated virus 6.
XX
PN WO2003104392-A2.
XX
PD 18-DEC-2003.
XX
XX 02-DEC-2002; 2002WO-US038423.
XX
XX 18-DEC-2001; 2001US-0341919P.
XX
XX (UYNC-) UNIV NORTH CAROLINA.
XX
XX Samulski RJ, Rabinowitz JE;
XX
XX WPI; 2004-062324/06.
XX
XX New polynucleotides comprising parvovirus rep coding sequences and
PT parvovirus cap coding sequences, useful in producing higher stocks of
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a
PT subject.
XX
XX Disclosure; SEQ ID NO 25; 115pp; English.
XX
XX The present invention describes a polynucleotide (I), comprising
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The
CC rep coding sequences encodes a DNA binding domain from a first
CC parvovirus, and a capsid interacting domain from a parvovirus different
CC from the first parvovirus. The cap coding sequence comprises sequences
CC from the different parvovirus. Also described: (1) a vector comprising
CC (1); (2) a cell comprising (1), or parvovirus rep coding sequences and
CC parvovirus cap coding sequences, where the rep coding sequences encode a
CC DNA binding domain from a first parvovirus and a capsid interacting
CC domain from a parvovirus different from the first parvovirus, the cap
CC coding sequences comprise sequences from the different parvovirus, and
CC the rep coding sequences are stably integrated into the genome of the
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in
CC gene therapy. The polynucleotide (I) can be used in producing higher
CC the delivery of nucleic acids having biological effect to treat or
CC ameliorate the symptoms associated with any disorder related to gene
CC expression. The polynucleotide may be used to produce a parvovirus vector
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.
CC The parvovirus vector may also be used to provide an antisense nucleic
CC acid to a cell in vitro or in vivo, or in diagnostic and screening
CC methods. The present sequence is used in the exemplification of the
CC present invention.
XX
XX Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 U; 0 Other;
SQ
Query Match 97.1%; Score 138.8; DB 12; Length 4683;
Best Local Similarity 98.6%; Pred. No. 3.5e-27;
Matches 140; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TTGCCACCTCCCTCTGCGCGCTCGCTCGTGGGCGCTGCGGACCAAGGTCCGC 60
Db 4683 TTGCCACCTCCCTCTGCGCGCTCGCTCGTGGGCGCTGCGGACCAAGGTCCGC 4624
QY 61 AGACGGGAGAGCTCTGCTGCGGCGCCACCCAGGAGCGGAGCGGAGGAGTG 120
Db 4623 AGACGGGAGAGCTCTGCTGCGGCGCCACCCAGGAGCGGAGCGGAGGAGTG 4564
QY 121 GGCAACTCCATCACTAGGGGTA 142
Db 4563 GGCAACTCCATCACTAGGGGTA 4542
```

## RESULT 9

ADE76502

XX ADE76502 standard; DNA; 4721 BP.

XX AC ADE76502;

XX DT 29-JAN-2004 (first entry)

XX DE

XX Adeno-associated virus (AAV) related DNA, SEQ ID No 1.

XX adeno-associated virus; AAV; cytotatic; antipsoriatic; antirheumatic;  
XX antiarthritic; neuroprotective; antidiabetic; antithyroid;  
XX dermatological; antinflammatory; gene therapy; vaccine;  
XX hyperproliferative; cancer; psoriasis; autoimmune disease;  
XX rheumatoid arthritis; multiple sclerosis; diabetes;  
XX autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.

XX OS

XX Adeno-associated virus 7.

XX PN

XX EP1310571-A2.

XX PD

XX 14-MAY-2003.

XX PF

XX 12-NOV-2002; 2002EP-00257826.

XX PR

XX 13-NOV-2001; 2001US-0350607P.

XX PR

XX 17-DEC-2001; 2001US-0341117P.

XX PR

XX 01-MAY-2002; 2002US-0377066P.

XX PR

XX 05-JUN-2002; 2002US-0386675P.

XX XX

(UYPE-) UNIV PENNSYLVANIA.

XX PA

XX Gao G, Wilson JM, Alvira M;

XX PI

XX WPI; 2003-450984/43.

XX DR

XX P-PSDB; ADE76503.

XX XX

XX Detecting adeno-associated virus sequences in a sample, useful for e.g.

XX PT

XX preventing or treating hyperproliferative or autoimmune diseases,

XX PT

XX comprises subjecting a sample having a DNA to amplification via

XX PT

XX polymerase chain reaction.

XX XX

XX Claim 17; SEQ ID NO 1; 419pp; English.

XX XX

XX The invention relates to a novel method for detecting adeno-associated  
XX virus (AAV) sequences in a sample, which comprises subjecting a sample  
XX containing a DNA to amplification via a polymerase chain reaction (PCR).  
XX The AAV sequence have the following activities: cytotatic,  
XX antipsoriatic, antirheumatic, antiarthritic, neuroprotective,  
XX antidiabetic, antithyroid, dermatological, and antinflammatory. The AAV  
XX sequence can be used in gene therapy or as part of a vaccine to treat  
XX disorders. The method is useful in detecting and/or identifying AAV  
XX sequences and isolating novel sequences that are identified. The  
XX sequences may be used e.g. for preventing or treating hyperproliferative  
XX conditions such as cancers and psoriasis, and other autoimmune diseases  
XX like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune  
XX thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence  
XX represents an AAV related DNA sequence of the invention.

XX SQ

XX Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;

XX

XX Query Match

XX Best Local Similarity 94.8%; Score 135.6; DB 10; Length 4721;

XX Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY

1 TTGCCACCTCCCTCTGCGCGCTCGCTCGTGGGCGCTGCGGACCAAGGTCCGC 60

Db

1 TTGGCCACTCCCTCTATGCGCGCTCGCTCGTGGGCGCTGCGGACCAAGGTCCGC 60

QY

61 AGACGGGAGAGCTCTGCTGCGGCGCCACCCAGGAGCGGAGCGGAGGAGTG 120

Db 61 AGACGGCAGAGCTCTGCTCTGCGGCCGCCACCGAGCGAGCGGCATAGAGGAGTG 120  
QY 121 GGCAACTCCATCACTAGGGGTA 142  
Db 121 GGCAACTCCATCACTAGGGGTA 142  
RESULT 10  
ADG39764  
ID ADG39764 standard; DNA; 4721 BP.  
XX  
AC ADG39764;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE AAV-7 genomic DNA sequence SEQ ID NO:26.  
XX  
KW parvovirus; rep; cap; DNA binding domain; capsid interacting domain;  
KW recombinant hybrid parvovirus particle;  
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;  
KW gene; ds.  
XX  
OS Adeno-associated virus 7.  
XX  
PN WO2003104392-A2.  
XX  
PD 18-DEC-2003.  
XX  
PF 02-DEC-2002; 2002WO-US038423.  
XX  
PR 18-DEC-2001; 2001US-0341919P.  
XX  
PA (UYNC-) UNIV NORTH CAROLINA.  
XX  
PI Samulski RJ, Rabinowitz JE;  
XX  
DR WPI; 2004-062324/06.  
XX  
XX New polynucleotides comprising parvovirus rep coding sequences and  
PT parvovirus cap coding sequences, useful in producing higher stocks of  
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a  
PT subject.  
XX  
PS Disclosure; SEQ ID NO 26; 115pp; English.  
XX  
CC The present invention describes a polynucleotide (I), comprising  
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The  
CC rep coding sequences encodes a DNA binding domain from a first  
CC parvovirus, and a capsid interacting domain from a parvovirus different  
CC from the first parvovirus. The cap coding sequence comprises sequences  
CC from the different parvovirus. Also described: (1) a vector comprising  
CC (1); (2) a cell comprising (1), or parvovirus rep coding sequences and  
CC parvovirus cap coding sequences, where the rep coding sequences encode a  
CC DNA binding domain from a first parvovirus and a capsid interacting  
CC domain from a parvovirus different from the first parvovirus, the cap  
CC coding sequences comprise sequences from the different parvovirus, and  
CC the rep coding sequences are stably integrated into the genome of the  
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno  
CC -associated virus (rAAV) particle. (1) can be used in vaccines, and in  
CC gene therapy. The polynucleotide (I) can be used in producing higher  
CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in  
CC the delivery of nucleic acids having biological effect to treat or  
CC ameliorate the symptoms associated with any disorder related to gene  
CC expression. The polynucleotide may be used to produce a parvovirus vector  
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.  
CC The parvovirus vector may also be used to provide an antisense nucleic  
CC acid to a cell in vitro or in vivo, or in diagnostic and screening  
CC methods. The present sequence is used in the exemplification of the  
CC present invention.  
XX  
SQ Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;  
Query Match 94.8%; Score 135.6; DB 12; Length 4721;

Best Local Similarity .97.2%; Pred. No. 2.5e-26;  
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 TTGCGCACCTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCCTGCGGACCAAGGTCCGC 60  
Db 1 TTGCGCACCTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCCTGCGGACCAAGGTCCGC 60  
QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCGCCACCGAGCGAGCGGCATAGAGGAGTG 120  
Db 61 AGACGGCAGAGCTCTGCTCTGCGGCCGCCACCGAGCGAGCGGCATAGAGGAGTG 120  
QY 121 GGCAACTCCATCACTAGGGGTA 142  
Db 121 GGCAACTCCATCACTAGGGGTA 142  
RESULT 11  
ADV67509  
ID ADV67509 standard; DNA; 4721 BP.  
XX  
AC ADV67509;  
XX  
DT 10-MAR-2005 (first entry)  
XX  
DE Nucleotide sequence of AAV serotype 7.  
XX  
KW antiarteriosclerotic; antilipemic; gene therapy; cholesterol;  
KW apolipoprotein E; apoE; apolipoprotein A; apoA; atherosclerosis;  
KW lipoprotein defect; ss.  
XX  
OS Adeno-associated virus.  
XX  
FH Key Location/Qualifiers  
FT repeat\_region 1..107  
FT /\*tag= a ITR"  
FT /\*note= "5' ITR"  
FT misc\_feature 825..3049  
FT /\*tag= b  
FT /\*note= "encodes vp1"  
FT misc\_feature 1234..3049  
FT /\*tag= c  
FT /\*note= "encodes vp2"  
FT misc\_feature 1434..3049  
FT /\*tag= d  
FT /\*note= "encodes vp3"  
FT repeat\_region 4704..4721  
FT /\*tag= e  
FT /\*note= "3' ITR"  
XX  
WO2004108922-A2.  
XX  
PD 16-DEC-2004.  
XX  
PF 23-APR-2004; 2004WO-US010965.  
XX  
PR 25-APR-2003; 2003US-0465293P.  
XX  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX  
PA Rader DJ, Wilson JM;  
XX WPI; 2005-031700/03.  
XX  
XX Lowering total cholesterol levels and treating atherosclerosis in a  
PT subject comprises delivering a recombinant adeno-associated virus (AAV)  
PT comprising an AAV serotype capsid protein or a gene encoding human  
PT apolipoprotein E (apoE) or apoA.  
XX  
PS Disclosure; SEQ ID NO 7; 69pp; English.  
XX  
CC The specification describes a method for lowering total cholesterol  
CC levels in a subject. The method comprises delivering to the subject a  
CC recombinant adeno-associated virus (AAV) comprising a gene encoding a

CC human apolipoprotein B (apoB) or apoA under the control of a regulatory  
 CC control sequence which directs expression of the gene. The recombinant  
 CC AAV also comprises a capsid protein selected from an AAV serotype (e.g.  
 CC AAV7 or AAV8) which preferentially expresses high levels of transgene in  
 CC live. A therapeutically effective amount of apoE or apoA expression is  
 CC obtained upon delivery of low dose of AAV. The method of the invention is  
 CC useful for lowering total cholesterol levels in a subject, e.g. for  
 CC treating atherosclerosis, and for correcting defects in lipoprotein. The  
 CC present sequence represents AAV serotype 7 (AAV7), which may be used to  
 CC produce recombinant AAV vectors of the invention.

XX  
 SQ Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;

Query Match 94.8%; Score 135.6; DB 14; Length 4721;  
 Best Local Similarity 97.2%; Pred. No. 2.5e-26;  
 Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TTGGCCACCTCCCTCTCTGCGCGCTCGCTCGGTCGGTGGGGCTCGCGACCAAGGTCCGC 60  
 DB 1 TTGGCCACCTCCCTCTCTATGCGCGCTCGCTCGGTCGGTGGGGCTCGCGACCAAGGTCCGC 60  
 QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGCGCAGAGAGGAGTG 120  
 DB 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGCGCAGAGGAGTG 120  
 QY 121 GGCACCTCCATCACTAGGGGTA 142  
 DB 121 GGCACCTCCATCACTAGGGGTA 142

## RESULT 12

ADZ27030  
 ID ADZ27030 standard; DNA; 4721 BP.

XX  
 AC ADZ27030;

XX 30-JUN-2005 (first entry)

XX Adeno-associated virus DNA SEQ ID NO 180.

XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;  
 KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;  
 KW bacterial infection; cancer; ulcerative colitis; antirheumatic;  
 KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;  
 KW antiapoptotic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;  
 KW virucide; antibacterial; cytostatic; antitumor; dermatological; ds.

XX Adeno-associated virus.

XX WO200503321-A2.

XX 14-APR-2005.

XX 30-SEP-2004; 2004WO-US028817.

XX 30-SEP-2003; 2003US-0508226P.

XX 29-APR-2004; 2004US-0566546P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;

XX WPI; 2005-285437/29.

XX New adeno-associated virus (AAV) clade comprising at least three AAV  
 PT members, useful for preventing and/or treating arthritis, multiple  
 PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial  
 PT infection and cancer.

XX Claim 12; SEQ ID NO 180; 569pp; English.

XX The invention relates to an adeno-associated virus (AAV) clade comprising  
 CC at least three AAV members, where each member of the AAV clade is

CC phylogenetically related as determined using a Neighbor-Joining heuristic  
 CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson  
 CC correction distance measurement of no more than 0.05. The methods and  
 CC compositions of the present invention are useful for the prevention  
 CC and/or treatment of rheumatoid arthritis, multiple sclerosis,  
 CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's  
 CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative  
 CC colitis. The present sequence represents an adeno-associated virus DNA.

XX  
 SQ Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;

Query Match 94.8%; Score 135.6; DB 14; Length 4721;  
 Best Local Similarity 97.2%; Pred. No. 2.5e-26;  
 Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TTGGCCACCTCCCTCTCTGCGCGCTCGCTCGGTCGGTGGGGCTCGCGACCAAGGTCCGC 60  
 DB 1 TTGGCCACCTCCCTCTCTATGCGCGCTCGCTCGGTCGGTGGGGCTCGCGACCAAGGTCCGC 60  
 QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGCGCAGAGAGGAGTG 120  
 DB 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCCACCAGCGAGCGCGCAGAGGAGTG 120  
 QY 121 GGCACCTCCATCACTAGGGGTA 142  
 DB 121 GGCACCTCCATCACTAGGGGTA 142

## RESULT 13

AAD00772/c

ID AAD00772 standard; DNA; 4718 BP.

XX AAD00772;

XX 08-SEP-2000 (first entry)

XX Adeno-associated virus serotype 1 DNA.

XX Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;  
 KW cap protein; recombinant viral vector; gene delivery; gene therapy;  
 KW vaccine; transgene; ss.

XX Adeno associated virus serotype 1.

FT Key Location/Qualifiers  
 FT repeat\_unit 1..143  
 FT /tag= a  
 FT /label= 5' ITR  
 FT /note= "Inverted terminal repeat which is capable of  
 FT forming T-shaped hairpin structure"

protein\_bind

FT 89..110  
 FT /tag= b

FT /bound\_moiety= "Rep protein"

misc\_feature

FT 124..125  
 FT /tag= c

FT /note= "Terminal resolve site (TRS)"

misc\_binding

FT 219..226  
 FT /tag= d

FT /bound\_moiety= "USF"

promoter

FT 236..299  
 FT /tag= e

FT /label= P5\_promoter

protein\_bind

FT 237..245  
 FT /tag= f

FT /bound\_moiety= "YY1 factor"

TATA\_signal

FT 270..275  
 FT /tag= g

FT /label= P5\_TATA-Box

misc\_feature

FT 299..306  
 FT /tag= h

FT /note= "YY1/p5 RNA"

FT CDS

FT 335..2272

```
FT      /*tag= j
FT      /product= "Rep 68"
FT      /function= "regulates replication and integration of AAV
FT      DNA into host cell's chromosome"
FT      /note= "The coding region is interrupted by intron"
FT      CDS
FT      335..2206
FT      /*tag= i
FT      /product= "Rep 78"
FT      /function= "regulates replication and integration of AAV
FT      DNA into host cell's chromosome"
FT      857..862
FT      TATA_signal
FT      /tag= l
FT      /label= P19_TATA_Box
FT      882..883
FT      /*tag= m
FT      /note= "P19 RNA"
FT      1007..2272
FT      CDS
FT      /*tag= o
FT      /product= "Rep 40"
FT      /function= "regulates replication and integration of AAV
FT      DNA into host cell's chromosome"
FT      /note= "The coding region is interrupted by intron"
FT      1007..2206
FT      /*tag= n
FT      /product= "Rep 52"
FT      /function= "regulates replication and integration of AAV
FT      DNA into host cell's chromosome"
FT      1836..1841
FT      TATA_signal
FT      /tag= p
FT      /label= P40_TATA-BOX
FT      1875..1876
FT      /*tag= q
FT      /note= "P40 RNA"
FT      1924..2220
FT      intron
FT      /tag= k
FT      /note= "This region interrupts the coding sequence of Rep
FT      68 and Rep 40"
FT      2223..4433
FT      CDS
FT      /*tag= r
FT      /product= "VP1 protein"
FT      /note= "Capsid protein"
FT      2634..4433
FT      CDS
FT      /*tag= s
FT      /product= "VP2 protein"
FT      /note= "Capsid protein"
FT      /partial
FT      2829..4433
FT      CDS
FT      /*tag= t
FT      /product= "VP3 protein"
FT      /note= "Capsid protein"
FT      4447..4452
FT      polyA_signal
FT      /tag= u
FT      4576..4718
FT      repeat_unit
FT      /*tag= v
FT      /label= 3' ITR
FT      /note= "Inverted terminal repeat which is capable of
FT      forming I-shaped hairpin structure"
FT      WO200028061-A2.
FT      PN
FT      18-MAY-2000.
FT      PD
FT      XX
FT      02-NOV-1999; 99WO-US025694.
FT      PF
FT      05-NOV-1998; 98US-0107114P.
FT      PR
FT      (UYPE-) UNIV PENNSYLVANIA.
FT      PA
FT      Wilson JM, Xiao W;
FT      PI
FT      WPI; 2000-376571/32.
FT      DR
FT      P-PSDB; AAY71161, AAY71164, AAY71165, AAY71166, AAY71167, AAY71168,
FT      DR      AAY71169.
```

```
XX      Novel adeno-associated virus serotype 1 polynucleotide useful for
PT      preparation of medicament for delivery of a transgene to a host.
XX      PS
XX      Claim 1; Fig 1; 108pp; English.
XX      CC
XX      The present sequence is an adeno-associated virus serotype 1 (AAV-1) DNA
CC      characterised by two inverted terminal repeats (ITR) and open reading
CC      frames for rep and capsid (cap) proteins. The rep reading frame encodes
CC      four proteins; Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading
CC      frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1
CC      sequence or its fragments particularly ITRs, rep and cap coding regions,
CC      are useful in production of recombinant viral vectors for gene delivery.
CC      These vectors can be used as gene therapy vectors, vaccine vectors or
CC      antisense delivery vectors. The AAV-1 does not induce the formation of
CC      neutralising antibodies specific to any serotype of AAV hence is useful
CC      for transforming host cells, and in the preparation of a medicament for
CC      the delivery of transgene to a host.
XX      SQ
XX      Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
SQ      Query Match      83.2%; Score 119; DB 3; Length 4718;
SQ      Best Local Similarity 89.5%; Pred. No. 5.8e-22;
SQ      Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY      1 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTGGTGGGGCTTGCAGACCAAGGTCGCG 60
DB      4718 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTGGTGGGGCTTGCAGACGAGCTCTGC 4659
QY      61 AGACGGCAGAGCTCTGCTCTGCGGCGCCCGCCAGCGAGCGAGCGCGAGAGGAGTG 120
DB      4658 CGTCTCGGACCTTTGGTCCGAGGCCCGCCAGCGAGCGAGCGAGCGAGAGGAGTG 4599
QY      121 GGCACCTCCATCCTAGGGGTAA 143
DB      4598 GGCACCTCCATCCTAGGGGTAA 4576
RESULT 14
ADE76507/C
ID      ADE76507 standard; DNA; 4718 BP.
XX
AC      ADE76507;
XX
DT      29-JAN-2004 (first entry)
XX
DE      Adeno-associated virus (AAV) related DNA, SEQ ID No 6.
XX
KW      adeno-associated virus; AAV; cytostatic; antipsoriatic; antirheumatic;
KW      antiarthritic; neuroprotective; antidiabetic; antithyroid;
KW      dermatological; antiinflammatory; gene therapy; vaccine;
KW      hyperproliferative; cancer; psoriasis; autoimmune disease;
KW      rheumatoid arthritis; multiple sclerosis; diabetes;
KW      autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.
XX
OS      Adeno-associated virus 1.
XX
XX      EP1310571-A2.
XX
XX      14-MAY-2003.
XX
XX      12-NOV-2002; 2002BP-00257826.
XX
XX      13-NOV-2001; 2001US-0350607P.
XX      17-DEC-2001; 2001US-0341117P.
XX      01-MAY-2002; 2002US-0377066P.
XX      05-JUN-2002; 2002US-0386675P.
XX
XX      (UYPE-) UNIV PENNSYLVANIA.
XX      PA
XX      Gao G, Wilson JM, Alvira M;
XX      WPI; 2003-450984/43.
XX      DR
```

XX Detecting adeno-associated virus sequences in a sample, useful for e.g.  
PT preventing or treating hyperproliferative or autoimmune diseases,  
PT comprises subjecting a sample having a DNA to amplification via  
PT polymerase chain reaction.  
XX  
XX  
PS Claim 14; SEQ ID NO 6; 419pp; English.  
PS  
XX The invention relates to a novel method for detecting adeno-associated  
CC virus (AAV) sequences in a sample, which comprises subjecting a sample  
CC containing a DNA to amplification via a polymerase chain reaction (PCR).  
CC The AAV sequence have the following activities: cytostatic,  
CC antiproliferative, antineoplastic, antitumor, neuroprotective,  
CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV  
CC sequences can be used in gene therapy or as part of a vaccine to treat  
CC disorders. The method is useful in detecting and/or identifying AAV  
CC sequences and isolating novel sequences that are identified. The  
CC sequences may be used e.g. for preventing or treating hyperproliferative  
CC conditions such as cancers and psoriasis, and other autoimmune diseases  
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune  
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence  
CC represents an AAV related DNA sequence of the invention.  
XX  
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;  
Query Match 83.2%; Score 119; DB 10; Length 4718;  
Best Local Similarity 89.5%; Pred. No. 5.8e-22;  
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 1 TTGCCCACCTCTCTGCGCGCTCGCTCGTGGGGCTGCGAGCAAGGTCCGC 60  
DB 4718 TTGCCCACCTCTCTGCGCGCTCGCTCGTGGGGCTGCGAGCAAGGTCTGC 4659  
QY 61 AGACGGCAGAGCTCTGCTTGGCGGCGCCACCGAGCGAGCGCGAGAGGGAGTG 120  
DB 4658 CGTCTGGGACCTTTGGTCCGACAGGCGCCACCGAGCGAGCGCGAGAGGGAGTG 4599  
QY 121 GGCAACTCCATCACTAGGGGTAA 143  
DB 4598 GGCAACTCCATCACTAGGGGTAA 4576  
RESULT 15  
ADL13984/C  
ID ADL13984 standard; DNA; 4718 BP.  
XX  
XX ADL13984;  
XX  
XX 06-MAY-2004 (first entry)  
XX  
XX Adeno-associated virus serotype 1 complete DNA.  
XX  
XX ss; cytostatic; neuroprotective; antiinflammatory; gene therapy;  
XX expression construct; adeno-associated virus;  
XX integration efficiency element; inverted terminal repeat; integration;  
XX chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;  
XX retinoblastoma; inflammatory disease; arthritis;  
XX neurodegenerative disease.  
XX  
XX Adeno-associated virus 1.  
XX  
XX WO2003087334-A2.  
XX  
XX 23-OCT-2003.  
XX  
XX 09-APR-2003; 2003WO-US011191.  
XX  
XX 09-APR-2002; 2002US-0371044P.  
XX  
XX (CORR ) CORNELL RES FOUND INC.  
XX  
XX Falck-Pedersen ES, Philpott N;  
XX  
XX

DR WPI; 2003-833723/77.  
XX  
XX New expression construct comprising a nucleic acid sequence encoding an  
PT adeno-associated virus integration efficiency element, useful for  
PT treating cancer e.g. lung cancer or colon cancer or inflammatory disease  
PT e.g. arthritis.  
XX  
XX Disclosure; SEQ ID NO 3; 62pp; English.  
XX  
XX The invention relates to an expression construct comprising a nucleic  
CC acid sequence encoding an adeno-associated virus integration efficiency  
CC element (AAV IEE), which is devoid of AAV inverted terminal repeats (AAV  
CC ITRs) and site-specifically integrates into a host cell chromosome when  
CC provided to the host cell in conjunction with an AAV Rep protein. The  
CC expression construct can be used as a therapeutic factor for treating a  
CC mammal for a pathologic state which is cancer, including lung cancer,  
CC colon cancer, renal cancer, anal cancer, bile duct cancer, bladder  
CC cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,  
CC cervical cancer, lymphoma, endometrial cancer, esophageal cancer,  
CC gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,  
CC liver cancer, multiple myeloma, neuroblastoma, skin cancer, stomach cancer,  
CC cancer, prostatic cancer, retinoblastoma, skin cancer, ovarian cancer, pancreatic  
CC testicular cancer, thymus cancer or thyroid cancer. Other pathologic  
CC state includes inflammatory disease (arthritis), neurodegenerative  
CC disease, a disease of an organ attributed to the presence of increased or  
CC decreased level of a particular gene product(s). This sequence  
CC corresponds to the AAV serotype 1 complete DNA including the IBE  
CC sequence.  
XX  
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;  
Query Match 83.2%; Score 119; DB 10; Length 4718;  
Best Local Similarity 89.5%; Pred. No. 5.8e-22;  
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 1 TTGCCCACCTCTCTGCGCGCTCGCTCGTGGGGCTGCGAGCAAGGTCCGC 60  
DB 4718 TTGCCCACCTCTCTGCGCGCTCGCTCGTGGGGCTGCGAGCAAGGTCTGC 4659  
QY 61 AGACGGCAGAGCTCTGCTTGGCGGCGCCACCGAGCGAGCGCGAGAGGGAGTG 120  
DB 4658 CGTCTGGGACCTTTGGTCCGACAGGCGCCACCGAGCGAGCGCGAGAGGGAGTG 4599  
QY 121 GGCAACTCCATCACTAGGGGTAA 143  
DB 4598 GGCAACTCCATCACTAGGGGTAA 4576

Search completed: November 28, 2005, 19:55:37

Job time : 69.9124 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 18:24:29 ; Search time 454.049 Seconds  
(without alignments)  
14735.325 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_1\_143  
Perfect score: 143  
Sequence: 1 tftgccactccctctctgcg.....aactccatcactaggggtaa 143

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	38.2	26.7	763	2	BG490707
c 2	37.8	26.4	1021	5	B0838990
c 3	37.4	26.2	644	2	BB632576
c 4	36.2	25.3	920	2	BG764368
c 5	36.2	25.3	932	10	CNS0072Q
c 6	36.2	25.3	1021	5	B0838990
c 7	35.8	25.0	604	3	B1531460
c 8	35.8	25.0	1074	10	CL507458
c 9	35.6	24.9	578	2	BB651469
c 10	35.4	24.8	860	2	BF578052
c 11	35.4	24.8	1123	10	AG080476
c 12	35.2	24.6	2024	10	AG074763
c 13	35	24.5	724	10	CL751558
c 14	35	24.5	844	10	CNS0052P
c 15	35	24.5	1303	9	BZ570288
c 16	34.8	24.3	676	9	BZ414423
c 17	34.8	24.3	926	8	CV775269
c 18	34.6	24.2	467	9	CC557657
c 19	34.6	24.2	617	7	CN038418
c 20	34.6	24.2	637	7	CN045740
c 21	34.6	24.2	2088	10	CL507788
c 22	34.4	24.1	318	7	CV352119

c 23	34.4	24.1	745	2	BG837739
c 24	34.4	24.1	890	10	AG127787
c 25	34.2	23.9	364	5	BY014749
c 26	34.2	23.9	468	8	CX358645
c 27	34.2	23.9	1367	10	CW943432
c 28	34	23.8	343	5	BY181775
c 29	34	23.8	450	5	BY250974
c 30	34	23.8	623	2	BB613613
c 31	34	23.8	631	5	BY723537
c 32	34	23.8	639	2	BB651977
c 33	34	23.8	643	2	BB652152
c 34	34	23.8	657	2	BB651054
c 35	34	23.8	661	2	BB621081
c 36	34	23.8	678	5	BY726305
c 37	34	23.8	795	9	AQ331299
c 38	34	23.8	838	8	CV780024
c 39	34	23.8	1364	5	BU543298
c 40	34	23.8	1767	4	AK082810
c 41	34	23.8	2268	4	AK082478
c 42	34	23.8	2708	4	AK038836
c 43	34	23.8	3798	4	AK031553
c 44	34	23.8	3994	4	AK029139
c 45	33.8	23.6	395	10	CW754890

ALIGNMENTS

RESULT 1  
BG490707/c  
LOCUS  
DEFINITION  
602520211f1 NIH\_MGC\_18 Homo sapiens CDNA clone IMAGE:4638699 5',  
mRNA sequence.  
ACCESSION  
BG490707  
VERSION  
BG490707.1  
KEYWORDS  
GI:13452217  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
REFERENCE  
1 (bases 1 to 763)  
NIH-MGC http://img.nci.nih.gov/.  
AUTHORS  
National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE  
Unpublished (1999)  
JOURNAL  
Contact: Robert Strausberg, Ph.D.  
COMMENT  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: DCTD/DP/Genard  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM1401 row: e column: 04  
High quality sequence stop: 182.

FEATURES  
source

Location/Qualifiers  
1..763  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4638699"  
/lab\_host="DH10B (phage-resistant)"  
/notes="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."











Query Match	24.5%;	Score 35;	DB 10;	Length 724;
Best Local Similarity	55.3%;	Pred. No. 70;	Mismatches 55;	Indels 0;
Matches	68;	Conservative	0;	Gaps 0;
QY	1	TTGCCACTCCCTCTCTGTCGGCGCTCGCTCGCTCGGTGGGGCTTCGCGACCAAAAGGTCGCG	60	
DB	438	TTATCTCTCTCTACTCTCTCTCTCTCACCAGAGCGCTGCGGAGGCGGCGGCTCGG	497	
QY	61	AGACGGCAGAGCTCTGCTCTTGGCGGCCCCACGAGCGAGCGGCGCGCAGAGGAGGTG	120	
DB	498	CGACGAAGCGGCGCTCGCTGGCGGGCTACGGGGGAGCGCGGCTCAATTACGAGGAG	557	
QY	121	GGC 123		
DB	558	GGC 560		
RESULT 14				
CNS0052P				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
source				
ORIGIN				
Query Match	24.5%;	Score 35;	DB 10;	Length 844;
Best Local Similarity	21.5%;	Pred. No. 70;	Mismatches 61;	Indels 34;
Matches	26;	Conservative	0;	Gaps 0;
QY	3	GCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCGCTGCGGACCAAGGTCGCGAG	62	
DB	490	SMCCSMCCSVSSVCAVCSSGVSRSVAVCGSVGVGSSRGRAGRSRGGGSSVSGW	549	
QY	63	ACGGCAGAGCTCTGCTCTGCCGCCCCACCGAGCGGACCGCGCAGAGGAGGTGGG	122	



Db 550 SSSSSVGMGCACSSASSVSCBSSVASSVSSGBVSRGCRVCVGGVGGSRVSSCSGSS 609  
 QY 123 C 123  
 Db 610 S 610

## RESULT 15

BZ570288

LOCUS

DEFINITION

BZ570288 1303 bp DNA linear GSS 17-DEC-2002

mesh2\_1301.x1 mesh Pseudomonas aeruginosa genomic clone mesh2\_1301,

genomic survey sequence.

ACCESSION BZ570288

VERSION BZ570288

KEYWORDS BZ570288.1 GI:27205349

SOURCE GSS.

ORGANISM Pseudomonas aeruginosa

Pseudomonas aeruginosa

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

Pseudomonadaceae; Pseudomonas.

REFERENCE 1 (bases 1 to 1303)

AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,

Burns,J.L., Kaul,R. and Olsen,M.V.

TITLE Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol. (2002) In press

JOURNAL J. Bacteriol. (2002) In press

COMMENT Contact: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

Location/Qualifiers

1. .1303

/organism="Pseudomonas aeruginosa"

/mol\_type="genomic DNA"

/strain="MSH"

/db\_xref="taxon:287"

/clone="mesh2\_1301"

/clone\_lib="mesh"

/note="Environmental isolate. Whole genomic shotgun

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

library."

Search completed: November 29, 2005, 14:17:49

Job time : 458.049 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 08:46:14 ; Search time 66.9124 Seconds  
(without alignments)  
14243.261 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_4576\_4718

Perfect score: 143

Sequence: 1 ttacccttagtgatggagtt.....cgcagaggggagtgggcaa 143

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- N\_Geneseq\_21.\*
- 1: Geneseqn1980s.\*
  - 2: Geneseqn1990s.\*
  - 3: Geneseqn2000s.\*
  - 4: Geneseqn2001as.\*
  - 5: Geneseqn2001bs.\*
  - 6: Geneseqn2002as.\*
  - 7: Geneseqn2002bs.\*
  - 8: Geneseqn2003as.\*
  - 9: Geneseqn2003bs.\*
  - 10: Geneseqn2003cs.\*
  - 11: Geneseqn2003ds.\*
  - 12: Geneseqn2004as.\*
  - 13: Geneseqn2004bs.\*
  - 14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	100.0	4718	3	AAD00772
2	143	100.0	4718	10	ADG39763
3	143	100.0	4718	10	ADL13984
4	143	100.0	4718	12	ADG39758
5	135.6	94.8	4721	10	ADG39750
6	135.6	94.8	4721	12	ADG39764
7	135.6	94.8	4721	14	ADG39764
8	135.6	94.8	4721	14	ADG39764
9	125	87.4	4718	3	AAD00772
10	121.8	85.2	4683	4	AAF23749
11	121.8	85.2	4683	10	ADL13983
12	121.8	85.2	4683	12	ADG39763
13	119	83.2	4718	3	AAD00772
14	119	83.2	4718	10	ADG39763
15	119	83.2	4718	10	ADG39763
16	119	83.2	4718	10	ADG39763
17	119	83.2	4718	12	ADG39758
18	114.8	80.3	4683	4	AAF23749
19	114.8	80.3	4683	10	ADL13983

20	114.8	80.3	4683	12	ADG39763	Adg39763 AAV-6 gen
21	113.2	79.2	4726	10	ADG39763	Adg39763 AAV-6 gen
22	113.2	79.2	4726	10	ADL13985	Adl13985 Adeno-ass
23	113.2	79.2	4726	12	ADG39759	Adg39759 AAV-3A ge
24	111.6	78.0	4721	10	ADG39764	Adg39764 AAV-7 gen
25	111.6	78.0	4721	12	ADG39764	Adg39764 AAV-7 gen
26	111.6	78.0	4721	14	ADG39764	Adg39764 AAV-7 gen
27	111.6	78.0	4721	14	ADG39764	Adg39764 AAV-7 gen
28	107.4	75.1	144	9	ACF35877	Adf35877 AAV-2 rig
29	107.4	75.1	145	6	ABA02990	AbA02990 Adeno-ass
30	107.4	75.1	145	6	ABA02990	AbA02990 Adeno-ass
31	107.4	75.1	145	13	ADM93254	Adm93254 Inverted
32	107.4	75.1	165	2	AAQ66769	AAQ66769 Double-D
33	107.4	75.1	165	2	AAQ66769	AAQ66769 Double-D
34	107.4	75.1	165	4	AAH41481	Aah41481 Adeno-ass
35	107.4	75.1	165	8	ABV77279	Abv77279 Nucleotid
36	107.4	75.1	192	2	AAT03384	Aat03384 Right-han
37	107.4	75.1	207	6	AAH41481	Aah41481 Adeno-ass
38	107.4	75.1	272	3	AAZ47166	Aaz47166 Head-to-t
39	107.4	75.1	272	3	AAZ47166	Aaz47166 Head-to-t
40	107.4	75.1	4675	4	AAH41481	Aah41481 Adeno-ass
41	107.4	75.1	4675	4	AAH41481	Aah41481 Adeno-ass
42	107.4	75.1	4675	6	ABA02989	AbA02989 Adeno-ass
43	107.4	75.1	4675	6	ABS69879	Abs69879 Human ade
44	107.4	75.1	4675	6	ABS69880	Abs69880 Human ade
45	107.4	75.1	4675	10	ADE76508	Ade76508 Adeno-ass

ALIGNMENTS

RESULT 1

AAD00772

ID AAD00772 standard; DNA; 4718 BP.

XX AAD00772;

XX 08-SEP-2000 (first entry)

XX Adeno-associated virus serotype 1 DNA.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

XX Adeno-associated virus serotype 1.

```
FT FT /*tag= h
FT FT /note= "VP1/p5 RNA"
FT FT 335..2272
FT FT /tag= j
FT FT /product= "Rep 68"
FT FT /function= "regulates replication and integration of AAV
FT FT DNA into host cell's chromosome"
FT FT /note= "The coding region is interrupted by intron"
FT FT 335..2206
FT FT /*tag= i
FT FT /product= "Rep 78"
FT FT /function= "regulates replication and integration of AAV
FT FT DNA into host cell's chromosome"
FT FT 857..862
FT FT /tag= l
FT FT /label= P19_TATA_Box
FT FT 882..883
FT FT /*tag= m
FT FT /note= "P19 RNA"
FT FT 1007..2272
FT FT /*tag= o
FT FT /product= "Rep 40"
FT FT /function= "regulates replication and integration of AAV
FT FT DNA into host cell's chromosome"
FT FT /note= "The coding region is interrupted by intron"
FT FT 1007..2206
FT FT /*tag= n
FT FT /product= "Rep 52"
FT FT /function= "regulates replication and integration of AAV
FT FT DNA into host cell's chromosome"
FT FT 1836..1841
FT FT /*tag= p
FT FT /label= P40_TATA_BOX
FT FT 1875..1876
FT FT /*tag= q
FT FT /note= "P40 RNA"
FT FT 1924..2220
FT FT /*tag= k
FT FT /note= "This region interrupts the coding sequence of Rep
FT FT 68 and Rep 40"
FT FT 2223..4433
FT FT /*tag= r
FT FT /product= "VP1 protein"
FT FT /note= "Capsid protein"
FT FT 2634..4433
FT FT /*tag= s
FT FT /product= "VP2 protein"
FT FT /note= "Capsid protein"
FT FT /partial
FT FT 2829..4433
FT FT /*tag= t
FT FT /product= "VP3 protein"
FT FT /note= "Capsid protein"
FT FT 4447..4452
FT FT /tag= u
FT FT /tag= v
FT FT 4576..4718
FT FT /*tag= v
FT FT /label= 3'_ITR
FT FT /note= "Inverted terminal repeat which is capable of
FT FT forming T-shaped hairpin structure"
FT FT
XX XX WO200028061-A2.
XX XX
XX XX 18-MAY-2000.
XX XX
XX XX 02-NOV-1999; 99WO-US025694.
XX XX
XX XX 05-NOV-1998; 98US-0107114P.
XX XX
XX XX (UYPE-) UNIV PENNSYLVANIA.
XX XX
XX XX Wilson JM, Xiao W;
XX XX
```

```
DR WPI; 2000-376571/32.
DR P-PSDB; AAY71161, AAY71164, AAY71165, AAY71166, AAY71167, AAY71168,
DR AAY71169.
XX
XX Novel adeno-associated virus serotype 1 polynucleotide useful for
XX preparation of medicament for delivery of a transgene to a host.
XX
XX Claim 1; Fig 1; 108pp; English.
XX
XX The present sequence is an adeno-associated virus serotype 1 (AAV-1) DNA
XX characterised by two inverted terminal repeats (ITR) and open reading
XX frames for rep and capsid (cap) proteins. The rep reading frame encodes
XX four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading
XX frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1
XX sequence or its fragments particularly ITRs, rep and cap coding regions,
XX are useful in production of recombinant viral vectors for gene delivery.
XX These vectors can be used as gene therapy vectors, vaccine vectors or
XX antisense delivery vectors. The AAV-1 does not induce the formation of
XX neutralising antibodies specific to any serotype of AAV hence is useful
XX for transforming host cells, and in the preparation of a medicament for
XX the delivery of transgene to a host
XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 143; DB 3; Length 4718;
Best Local Similarity 100.0%; Pred. No. 3.6e-28;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGTGAGATTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCC 60
Db 4576 TTACCCCTAGTGTGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCC 4635
QY 61 TCGCGACCAAGGTCCGACAGCGGAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGA 120
Db 4636 TCGCGACCAAGGTCCGACAGCGGAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGA 4695
QY 121 GCGCGACAGAGGAGTGGGCAA 143
Db 4696 GCGCGACAGAGGAGTGGGCAA 4718
RESULT 2
ADE76507
ID ADE76507 standard; DNA; 4718 BP.
XX
XX ADE76507;
XX
XX 29-JAN-2004 (first entry)
XX
XX Adeno-associated virus (AAV) related DNA, SEQ ID No 6.
XX
XX adeno-associated virus; AAV; cytostatic; antipsoriatic; antirheumatic;
XX antiarthritic; neuroprotective; antidiabetic; antithyroid;
XX dermatological; antiinflammatory; gene therapy; vaccine;
XX hyperproliferative; cancer; psoriasis; autoimmune disease;
XX rheumatoid arthritis; multiple sclerosis; diabetes;
XX autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.
XX
XX Adeno-associated virus 1.
XX
XX EP1310571-A2.
XX
XX PD 14-MAY-2003.
XX
XX 12-NOV-2002; 2002EP-00257826.
XX
XX 13-NOV-2001; 2001US-0350607P.
XX
XX 17-DEC-2001; 2001US-0341117P.
XX
XX 01-MAY-2002; 2002US-0377066P.
XX
XX 05-JUN-2002; 2002US-0386675P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX
```

```
PI Gao G, Wilson JM, Alvira M;
XX WPI; 2003-450984/43.
XX
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.
PT preventing or treating hyperproliferative or autoimmune diseases,
PT comprises subjecting a sample having a DNA to amplification via
PT polymerase chain reaction.
XX
XX Claim 14; SEQ ID NO 6; 419pp; English.
XX
XX The invention relates to a novel method for detecting adeno-associated
CC virus (AAV) sequences in a sample, which comprises subjecting a sample
CC containing a DNA to amplification via a polymerase chain reaction (PCR).
CC The AAV sequence have the following activities: cytostatic,
CC antipsoriatic, antirheumatic, antilarthritic, neuroprotective,
CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
CC sequence can be used in gene therapy or as part of a vaccine to treat
CC disorders. The method is useful in detecting and/or identifying AAV
CC sequences and isolating novel sequences that are identified. The
CC sequences such as cancers and psoriasis, and other autoimmune diseases
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence
CC represents an AAV related DNA sequence of the invention.
XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
Query Match 100.0%; Score 143; DB 10; Length 4718;
Best Local Similarity 100.0%; Pred. No. 3.6e-28;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGAGGTGCCACTCTCTCTCGCGCTCGCTCGCTCGCTCGGTGGGGCC 60
DB 4576 TTACCCCTAGTGAGGTGCCACTCTCTCTCGCGCTCGCTCGCTCGCTCGGTGGGGCC 4635
QY 61 TCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTCGCGCCCCACCGAGCGAGCGA 120
DB 4636 TCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTCGCGCCCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGAGGAGTGGGCAA 143
DB 4696 GCGCGCAGAGAGGAGTGGGCAA 4718
RESULT 3
ADL13984
ID ADL13984 standard; DNA; 4718 BP.
XX
XX ADL13984;
XX
XX 06-MAY-2004 (first entry)
XX
XX Adeno-associated virus serotype 1 complete DNA.
DE
XX
XX sa; cytostatic; neuroprotective; antiinflammatory; gene therapy;
XX expression construct; adeno-associated virus;
XX integration efficiency element; inverted terminal repeat; integration;
XX chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;
XX retinoblastoma; inflammatory disease; arthritis;
XX neurodegenerative disease.
XX
XX Adeno-associated virus 1.
OS
XX WO2003087334-A2.
XX
XX 23-OCT-2003.
XX
XX 09-APR-2003; 2003WO-US011191.
XX
XX 09-APR-2002; 2002US-0371044P.
XX
XX (CORR ) CORNELL RES FOUND INC.
```

```
XX Falck-Pedersen ES, Philpott N;
PI
XX WPI; 2003-833723/77.
XX
XX New expression construct comprising a nucleic acid sequence encoding an
PT adeno-associated virus integration efficiency element, useful for
PT treating cancer e.g. lung cancer or colon cancer or inflammatory disease
PT e.g. arthritis.
XX
XX Disclosure; SEQ ID NO 3; 62pp; English.
XX
XX The invention relates to an expression construct comprising a nucleic
CC acid sequence encoding an adeno-associated virus integration efficiency
CC element (AAV IEE), which is devoid of AAV inverted terminal repeats (AAV
CC ITRs) and site-specifically integrates into a host cell chromosome when
CC provided to the host cell in conjunction with an AAV Rep protein. The
CC expression construct can be used as a therapeutic factor for treating a
CC mammal for a pathologic state which is cancer, including lung cancer,
CC colon cancer, renal cancer, anal cancer, bile duct cancer, bladder
CC cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,
CC cervical cancer, lymphoma, endometrial cancer, esophageal cancer,
CC gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,
CC liver cancer, multiple myeloma, neuroblastoma, ovarian cancer, pancreatic
CC cancer, prostatic cancer, retinoblastoma, skin cancer, stomach cancer,
CC testicular cancer, thymus cancer or thyroid cancer. Other pathologic
CC state includes inflammatory disease (arthritis), neurodegenerative
CC disease, a disease of an organ attributed to the presence of increased or
CC decreased level of a particular gene product(s). This sequence
CC corresponds to the AAV serotype 1 complete DNA including the IEE
CC sequence.
XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
Query Match 100.0%; Score 143; DB 10; Length 4718;
Best Local Similarity 100.0%; Pred. No. 3.6e-28;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGAGGTGCCACTCTCTCTCGCGCTCGCTCGCTCGGTGGGGCC 60
DB 4576 TTACCCCTAGTGAGGTGCCACTCTCTCTCGCGCTCGCTCGCTCGGTGGGGCC 4635
QY 61 TCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTCGCGCCCCACCGAGCGAGCGA 120
DB 4636 TCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTCGCGCCCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGAGGAGTGGGCAA 143
DB 4696 GCGCGCAGAGAGGAGTGGGCAA 4718
RESULT 4
ADG39758
ID ADG39758 standard; DNA; 4718 BP.
XX
XX ADG39758;
XX
XX 11-MAR-2004 (first entry)
XX
XX AAV-1 genomic DNA sequence SEQ ID NO:20.
DE
XX
XX parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
XX recombinant hybrid parvovirus particle;
XX recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
XX gene; ds.
XX
XX Adeno-associated virus 1.
OS
XX WO2003104392-A2.
XX
XX 18-DEC-2003.
XX
XX 02-DEC-2002; 2002WO-US038423.
```

XX 18-DEC-2001; 2001US-0341919P.  
XX (UYN-) UNIV NORTH CAROLINA.  
XX Samulski RJ, Rabinowitz JE;  
XX MPI; 2004-062324/06.  
XX  
XX New polynucleotides comprising parvovirus rep coding sequences and  
PT parvovirus cap coding sequences, useful in producing higher stocks of  
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a  
PT subject.  
XX  
XX Disclosure; SEQ ID NO 20; 115pp; English.  
XX  
XX The present invention describes a polynucleotide (I), comprising  
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The  
CC rep coding sequences encodes a DNA binding domain from a first  
CC parvovirus, and a capsid interacting domain from a parvovirus different  
CC from the first parvovirus. The cap coding sequence comprises sequences  
CC from the different parvovirus. Also described: (1) a vector comprising  
CC (I); (2) a cell comprising (I), or parvovirus rep coding sequences and  
CC parvovirus cap coding sequences, where the rep coding sequences encode a  
CC DNA binding domain from a first parvovirus and a capsid interacting  
CC domain from a parvovirus different from the first parvovirus, the cap  
CC coding sequences comprises sequences from the different parvovirus, and  
CC the rep coding sequences are stably integrated into the genome of the  
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno  
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in  
CC gene therapy. The polynucleotide (I) can be used in producing higher  
CC stocks of hybrid parvovirus or parvovirus vectors, which may be used in  
CC the delivery of nucleic acids having biological effect to treat or  
CC ameliorate the symptoms associated with any disorder related to gene  
CC expression. The polynucleotide may be used to produce a parvovirus vector  
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.  
CC The parvovirus vector may also be used to provide an antisense nucleic  
CC acid to a cell in vitro or in vivo, or in diagnostic and screening  
CC methods. The present sequence is used in the exemplification of the  
CC present invention.  
XX  
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 143; DB 12; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 3.6e-28;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCC 60  
DB 4576 TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCC 4635  
QY 61 TCGGACCAAAAGTCCGACAGCGCAGAGCTTGTCTCTGCCGGCCCCACCGAGCGAGA 120  
DB 4636 TCGGACCAAAAGTCCGACAGCGCAGAGCTTGTCTCTGCCGGCCCCACCGAGCGAGA 4695  
QY 121 GCGCGCAGAGGGAGTGGGCAA 143  
DB 4696 GCGCGCAGAGGGAGTGGGCAA 4718  
RESULT 5  
ADE76502  
ID ADE76502 standard; DNA; 4721 BP.  
XX  
XX ADE76502;  
AC  
XX 29-JAN-2004 (first entry)  
DT  
XX Adeno-associated virus (AAV) related DNA, SEQ ID No 1.  
DE  
XX adeno-associated virus; AAV; cytosolic; antipositive; antitumoratic;  
XX antitumoratic; antipositive; antidiabetic; antitumoratic;  
XX dermatological; antitumoratic; gene therapy; vaccine;  
KW

KW hyperproliferative; cancer; psoriasis; autoimmune disease;  
KW rheumatoid arthritis; multiple sclerosis; diabetes;  
XX autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.  
OS Adeno-associated virus 7.  
XX  
XX EPI310571-A2.  
XX  
XX 14-MAY-2003.  
XX  
XX 12-NOV-2002; 2002EP-00257826.  
XX  
XX 13-NOV-2001; 2001US-0350607P.  
XX 17-DEC-2001; 2001US-0341117P.  
XX 01-MAY-2002; 2002US-0377066P.  
XX 05-JUN-2002; 2002US-0386675P.  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX  
XX Gao G, Wilson JM, Alvira M;  
XX MPI; 2003-450984/43.  
XX P-PSDB; ADE76503.  
XX  
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.  
PT preventing or treating hyperproliferative or autoimmune diseases,  
PT comprises subjecting a sample having a DNA to amplification via  
PT polymerase chain reaction.  
XX  
XX Claim 17; SEQ ID NO 1; 419pp; English.  
XX  
XX The invention relates to a novel method for detecting adeno-associated  
CC virus (AAV) sequences in a sample, which comprises subjecting a sample  
CC containing a DNA to amplification via a polymerase chain reaction (PCR).  
CC The AAV sequence have the following activities: cytostatic,  
CC antipositive, antitumoratic, antidiabetic, antitumoratic, antitumoratic,  
CC antidiabetic, antitumoratic, dermatological, and antitumoratic. The AAV  
CC sequence can be used in gene therapy or as part of a vaccine to treat  
CC disorders. The method is useful in detecting and/or identifying AAV  
CC sequences and isolating novel sequences that are identified. The  
CC sequences may be used e.g. for preventing or treating hyperproliferative  
CC conditions such as cancer and psoriasis, and other autoimmune diseases  
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune  
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence  
CC represents an AAV related DNA sequence of the invention.  
XX  
XX Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;  
SQ  
Query Match 94.8%; Score 135.6; DB 10; Length 4721;  
Best Local Similarity 97.2%; Pred. No. 3.1e-26;  
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 2 TACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCC 61  
DB 4580 TACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCC 4639  
QY 62 GCGGACCAAAAGTCCGACAGCGCAGAGCTTGTCTCTGCCGGCCCCACCGAGCGAG 121  
DB 4640 GCGGACCAAAAGTCCGACAGCGCAGAGCTTGTCTCTGCCGGCCCCACCGAGCGAG 4699  
QY 122 GCGGACAGAGGGAGTGGGCAA 143  
DB 4700 GCGGACAGAGGGAGTGGGCAA 4721  
RESULT 6  
ADG39764  
ID ADG39764 standard; DNA; 4721 BP.  
XX  
XX ADG39764;  
AC  
XX 11-MAR-2004 (first entry)  
DT  
XX

DE AAV-7 genomic DNA sequence SEQ ID NO:26.  
XX parvovirus; rep; cap; DNA binding domain; capsid interacting domain;  
KW recombinant hybrid parvovirus particle;  
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;  
KW gene; ds.  
XX  
XX  
OS Adeno-associated virus 7.  
XX  
XX WO2003104392-A2.  
XX  
XX  
PD 18-DEC-2003.  
XX  
XX  
PF 02-DEC-2002; 2002WO-US038423.  
XX  
XX 18-DEC-2001; 2001US-0341919P.  
XX  
XX (UYNC-) UNIV NORTH CAROLINA.  
XX  
XX Samulski RJ, Rabinowitz JE;  
XX  
XX WPI; 2004-062324/06.  
XX  
XX New polynucleotides comprising parvovirus rep coding sequences and  
PT parvovirus cap coding sequences, useful in producing higher stocks of  
PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a  
PT subject.  
XX  
XX  
PS Disclosure; SEQ ID NO 26; 115pp; English.  
XX  
XX The present invention describes a polynucleotide (I), comprising  
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The  
CC rep coding sequences encodes a DNA binding domain from a first  
CC parvovirus, and a capsid interacting domain from a parvovirus different  
CC from the first parvovirus. The cap coding sequence comprises sequences  
CC from the different parvovirus. Also described: (1) a vector comprising  
CC (1); (2) a cell comprising (1), or parvovirus rep coding sequences and  
CC parvovirus cap coding sequences, where the rep coding sequences encode a  
CC DNA binding domain from a first parvovirus and a capsid interacting  
CC domain from a parvovirus different from the first parvovirus, the cap  
CC coding sequences comprise sequences from the different parvovirus, and  
CC the rep coding sequences are stably integrated into the genome of the  
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno  
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in  
CC gene therapy. The polynucleotide (I) can be used in producing higher  
CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in  
CC the delivery of nucleic acids having biological effect to treat or  
CC ameliorate the symptoms associated with any disorder related to gene  
CC expression. The polynucleotide may be used to produce a parvovirus vector  
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.  
CC The parvovirus vector may also be used to provide an antisense nucleic  
CC acid to a cell in vitro or in vivo, or in diagnostic and screening  
CC methods. The present sequence is used in the exemplification of the  
CC present invention.  
XX  
XX  
SQ Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;  
Query Match 94.8%; Score 135.6; DB 12; Length 4721;  
Best Local Similarity 97.2%; Pred. No. 3.1e-26;  
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 2 TACCCCTAGTGATGGAGTGGCCACATCCCTCTCTGCGCGCTCGCTCGTGGGGCCCT 61  
Db 4580 TACCCCTAGTGATGGAGTGGCCACATCCCTCTCTGCGCGCTCGCTCGTGGGGCCCT 4639  
QY 62 GCGGACCAAGGTCCGACAGCGCAGAGTCTCTCTGCGCGGCCACCGACGAGCGAG 121  
Db 4640 GCGGACCAAGGTCCGACAGCGCAGAGTCTCTCTGCGCGGCCACCGACGAGCGAG 4699  
QY 122 CGCGCAGAGAGGAGTGGGCAA 143  
Db 4700 CGCGCATAGAGGAGTGGGCAA 4721

RESULT 7  
ADV67509  
ID ADV67509 standard; DNA; 4721 BP.  
XX  
XX AC ADV67509;  
XX  
XX 10-MAR-2005 (first entry)  
XX  
XX Nucleotide sequence of AAV serotype 7.  
XX  
XX antiarteriosclerotic; antilipemic; gene therapy; cholesterol;  
KW apolipoprotein E; apoE; apolipoprotein A; apoA; atherosclerosis;  
KW lipoprotein defect; ss.  
XX  
XX Adeno-associated virus.  
XX  
XX  
FH Key Location/Qualifiers  
FT repeat\_region 1..107  
FT /\*tag= a  
FT /note= "5, ITR"  
FT misc\_feature 825..3049  
FT /\*tag= b  
FT /note= "encodes vp1"  
FT misc\_feature 1234..3049  
FT /\*tag= c  
FT /note= "encodes vp2"  
FT misc\_feature 1434..3049  
FT /\*tag= d  
FT /note= "encodes vp3"  
FT repeat\_region 4704..4721  
FT /\*tag= e  
FT /note= "3, ITR"  
XX  
XX WO2004108922-A2.  
XX  
XX 16-DEC-2004.  
XX  
XX 23-APR-2004; 2004WO-US010965.  
XX  
XX 25-APR-2003; 2003US-0465293P.  
XX  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX  
XX Rader DJ, Wilson JM;  
XX  
XX WPI; 2005-031700/03.  
XX  
XX Lowering total cholesterol levels and treating atherosclerosis in a  
PT subject comprises delivering a recombinant adeno-associated virus (AAV)  
PT comprising an AAV serotype capsid protein or a gene encoding human  
PT apolipoprotein E (apoE) or apoA.  
XX  
XX Disclosure; SEQ ID NO 7; 69pp; English.  
XX  
XX The specification describes a method for lowering total cholesterol  
CC levels in a subject. The method comprises delivering to the subject a  
CC recombinant adeno-associated virus (AAV) comprising a gene encoding a  
CC human apolipoprotein E (apoE) or apoA under the control of a regulatory  
CC control sequence which directs expression of the gene. The recombinant  
CC AAV also comprises a capsid protein selected from an AAV serotype (e.g.  
CC AAV7 or AAV8) which preferentially expresses high levels of transgene in  
CC live. A therapeutically effective amount of apoE or apoA expression is  
CC obtained upon delivery of low dose of AAV. The method of the invention is  
CC useful for lowering total cholesterol levels in a subject, e.g. for  
CC treating atherosclerosis, and for correcting defects in lipoprotein. The  
CC present sequence represents AAV serotype 7 (AAV7), which may be used to  
CC produce recombinant AAV vectors of the invention.  
XX  
XX  
SQ Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;  
Query Match 94.8%; Score 135.6; DB 14; Length 4721;  
Best Local Similarity 97.2%; Pred. No. 3.1e-26;  
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TACCCCTAGTGTGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCT 61  
 Db TACCCCTAGTGTGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCT 4639

QY 62 GCGGACCAAAAGTCCGACAGCGGAGAGCTCTGCTCTGCGGCTCGCTCGGTGGGCT 121  
 Db GCGGACCAAAAGTCCGACAGCGGAGAGCTCTGCTCTGCGGCTCGCTCGGTGGGCT 4699

QY 122 CGCGCAGAGAGGAGTGGGCAA 143  
 Db CGCGCAGAGAGGAGTGGGCAA 4721

## RESULT 8

ADZ27030

ID ADZ27030 standard; DNA; 4721 BP.

AC

ADZ27030;

XX

DT 30-JUN-2005 (first entry)

XX

DE Adeno-associated virus DNA SEQ ID NO 180.

XX

KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;  
 KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;  
 KW bacterial infection; cancer; ulcerative colitis; antirheumatic;  
 KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;  
 KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;  
 KW viricide; antibacterial; cytostatic; antiulcer; dermatological; ds.

XX

OS Adeno-associated virus.

XX

PN WO200503321-A2.

XX

PD 14-APR-2005.

XX

PF 30-SEP-2004; 2004WO-US028817.

XX

PR 30-SEP-2003; 2003US-0508226P.

XX

PR 29-APR-2004; 2004US-0566546P.

XX

PA (UYPE-) UNIV PENNSYLVANIA.

XX

PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;

XX

DR WPI; 2005-285437/29.

XX

PT New adeno-associated virus (AAV) clade comprising at least three AAV members, useful for preventing and/or treating arthritis, multiple sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial infection and cancer.

XX

PS Claim 12; SEQ ID NO 180; 569pp; English.

XX

CC The invention relates to an adeno-associated virus (AAV) clade comprising at least three AAV members, where each member of the AAV clade is phylogenetically related as determined using a Neighbor-Joining heuristic by a bootstrap value of at least 75 % per 1000 isolates and a Poisson correction distance measurement of no more than 0.05. The methods and compositions of the present invention are useful for the prevention and/or treatment of rheumatoid arthritis, multiple sclerosis, sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's disease, hemophilia, HIV, bacterial infection, cancer and ulcerative colitis. The present sequence represents an adeno-associated virus DNA.

XX

SQ Sequence 4721 BP; 1108 A; 1405 C; 1301 G; 907 T; 0 U; 0 Other;

XX

Query Match 94.8%; Score 135.6; DB 14; Length 4721;

Best Local Similarity 97.2%; Pred. No. 3.1e-26;

Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TACCCCTAGTGTGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCT 61  
 Db TACCCCTAGTGTGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCT 4639

QY 62 GCGGACCAAAAGTCCGACAGCGGAGAGCTCTGCTCTGCGGCTCGCTCGGTGGGCT 121  
 Db GCGGACCAAAAGTCCGACAGCGGAGAGCTCTGCTCTGCGGCTCGCTCGGTGGGCT 4699

QY 122 CGCGCAGAGAGGAGTGGGCAA 143  
 Db CGCGCAGAGAGGAGTGGGCAA 4721

## RESULT 9

AAD00781

ID AAD00781 standard; DNA; 143 BP.

XX

AC AAD00781;

XX

DT 08-SEP-2000 (first entry)

XX

DE Adeno-associated virus serotype 1 5' inverted terminal repeat.

XX

KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;  
 KW cap protein; recombinant viral vector; gene delivery; gene therapy;  
 KW vaccine; transgene; inverted terminal repeat; ITR; ss.

OS

Adeno associated virus serotype 1.

PH

Key Location/Qualifiers

FT misc\_binding

FT 1..41

FT /tag= a

FT /bound\_moiety= "binds to nucleotides 85..125"

FT stem\_loop

FT 42..62

FT /tag= b

FT stem\_loop

FT 64..84

FT /tag= c

FT misc\_binding

FT 85..125

FT /tag= d

FT /bound\_moiety= "binds to nucleotides 41..1"

XX

WO200028061-A2.

XX

PD 18-MAY-2000.

XX

PF 02-NOV-1999; 99WO-US025694.

XX

PR 05-NOV-1998; 98US-0107114P.

XX

(UYPE-) UNIV PENNSYLVANIA.

XX

PI Wilson JM, Xiao W;

XX

WPI; 2000-376571/32.

XX

Novel adeno-associated virus serotype 1 polynucleotide useful for

XX

preparation of medicament for delivery of a transgene to a host.

XX

Example 2; Fig 2; 108pp; English.

XX

CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA which is characterised by two inverted terminal repeats (ITR) and open reading frames for rep and capsid (cap) proteins. The rep reading frame encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1 sequence or its fragments particularly ITRs, rep and cap coding regions are useful in production of recombinant viral vectors for gene delivery. These vectors can be used as gene therapy vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does not induce the formation of neutralising antibodies specific to any serotype of AAV hence is useful for transforming host cells, and in the preparation of a medicament for the delivery of transgene to a host. The present sequence is a 5' ITR of AAV-1 DNA which is useful in the production of recombinant



CC viral vector. The ITR forms a T-shaped hairpin structure  
XX  
SQ Sequence 143 BP; 24 A; 50 C; 48 G; 21 T; 0 U; 0 Other;

Query Match 87.4%; Score 125; DB 3; Length 143;  
Best Local Similarity 100.0%; Pred. No. 1.6e-23;  
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTCGCGACCAAGGTCCGC 78  
DB 1 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTCGCGACCAAGGTCCGC 60

QY 79 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGGGAGTG 138  
DB 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGGGAGTG 120

QY 139 GGCNA 143  
DB 121 GGCNA 125

RESULT 10  
AAF23749/c  
ID AAF23749 standard; DNA; 4683 BP.

AC AAF23749;

DT 28-MAR-2001 (first entry)

XX AAV6 DNA sequence.

XX AAV6; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;  
KW atherosclerosis; sickle cell anaemia; thalassaemia;  
KW blood clotting disorder; diabetes; ss.

XX Adeno associated virus.

XX US6156303-A.

XX 05-DEC-2000.

XX 11-JUN-1997; 97US-00873168.

XX 11-JUN-1997; 97US-00873168.

XX (UNIW ) UNIV WASHINGTON.

XX Russell DW, Rutledge EA;

XX WPI; 2001-060164/07.

XX Adeno-associated virus serotype 6 and viral vector derived from it for  
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency  
PT syndrome, sickle cell anemia, thalassaemia and diabetes.

XX Claim 1; Fig 1; 50pp; English.

XX The present invention relates to adeno-associated virus serotypes. The  
CC present sequence is the DNA sequence of one such serotype (AAV6). AAV6  
CC can be used to construct AAV viral vectors for use in gene therapy for a  
CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,  
CC sickle cell anaemia, thalassaemia, blood clotting disorders and diabetes.  
CC The AAV viral vectors have increased transduction efficiency of a  
CC particular host cell as the AAV virion containing the AAV vector genome  
CC can be modified to express a capsid protein of an AAV serotype that  
CC transduces the selected host cell

XX Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 U; 0 Other;

Query Match 85.2%; Score 121.8; DB 4; Length 4683;  
Best Local Similarity 98.4%; Pred. No. 1.3e-22;  
Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 19 TTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTCGCGACCAAGGTCCGC 78  
DB 4683 TTGCCCACTCCCTCTATCGCGCTCGCTCGCTCGGTGGGGCTCGCGACCAAGGTCCGC 4624

QY 79 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGGGAGTG 138  
DB 4623 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGGGAGTG 4564

QY 139 GGCNA 143  
DB 4563 GGCNA 4559

ADL13983/c  
ID ADL13983 standard; DNA; 4683 BP.

XX ADL13983;

XX 06-MAY-2004 (first entry)

XX Adeno-associated virus serotype 6 complete DNA.

XX ss; cytostatic; neuroprotective; antiinflammatory; gene therapy;

XX expression construct; adeno-associated virus;

XX integration efficiency element; inverted terminal repeat; integration;

XX chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;

XX retinoblastoma; inflammatory disease; arthritis;

XX neurodegenerative disease.

XX Adeno-associated virus 6.

XX WO2003087334-A2.

XX 23-OCT-2003.

XX 09-APR-2003; 2003WO-US011191.

XX 09-APR-2002; 2002US-0371044P.

XX (CORR ) CORNELL RES FOUND INC.

XX Falck-Pedersen ES, Philpott N;

XX WPI; 2003-833723/77.

XX New expression construct comprising a nucleic acid sequence encoding an  
PT adeno-associated virus integration efficiency element, useful for  
PT treating cancer e.g. lung cancer or colon cancer or inflammatory disease  
PT e.g. arthritis.

XX Disclosure; SEQ ID NO 2; 62pp; English.

XX The invention relates to an expression construct comprising a nucleic  
CC acid sequence encoding an adeno-associated virus integration efficiency  
CC element (AAV IEE), which is devoid of AAV inverted terminal repeats (AAV  
CC ITRs) and site-specifically integrates into a host cell chromosome when  
CC provided to the host cell in conjunction with an AAV Rep protein. The  
CC expression construct can be used as a therapeutic factor for treating a  
CC mammal for a pathologic state which is cancer, including lung cancer,  
CC colon cancer, renal cancer, anal cancer, bile duct cancer, bladder  
CC cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,  
CC cervical cancer, lymphoma, endometrial cancer, esophageal cancer,  
CC gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,  
CC liver cancer, multiple myeloma, neuroblastoma, ovarian cancer, pancreatic  
CC cancer, prostatic cancer, retinoblastoma, skin cancer, stomach cancer,  
CC testicular cancer, thymus cancer or thyroid cancer. Other pathologic  
CC state includes inflammatory disease (arthritis), neurodegenerative  
CC disease, a disease of an organ attributed to the presence of increased or  
CC decreased level of a particular gene product(s). This sequence  
CC corresponds to the AAV serotype 6 complete DNA including the IEE  
CC sequence.

XX

```
SQ Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 U; 0 Other;
Query Match 85.2%; Score 121.8; DB 10; Length 4683;
Best Local Similarity 98.4%; Pred. No. 1.3e-22;
Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 19 TTGCCCACTCCTCTCTGCGGCTCGCTCGCTCGGTGGGGCTTGGGACCAAGGTCCGC 78
DB 4683 TTGCCCACTCCTCTATGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCAAGGTCCGC 4624
QY 79 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 138
DB 4623 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 4564
QY 139 GGCAA 143
DB 4563 GGCAA 4559
RESULT 13
ADG39763/C
ID ADG39763 standard; DNA; 4683 BP.
XX AC ADG39763;
XX DT 11-MAR-2004 (first entry)
XX DE AAV-6 genomic DNA sequence SEQ ID NO:25.
XX KW parvovirus; rep; cap; DNA binding domain; capsid interacting domain;
XX KW recombinant hybrid parvovirus particle;
XX KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;
XX KW Gene; db.
XX OS Adeno-associated virus 6.
XX PN WO2003104392-A2.
XX PD 18-DEC-2003.
XX PF 02-DEC-2002; 2002WO-US038423.
XX PR 18-DEC-2001; 2001US-0341919P.
XX PA (UYNC-) UNIV NORTH CAROLINA.
XX PI Samulski RJ, Rabinowitz JE;
XX DR WPI; 2004-062324/06.
XX PT New polynucleotides comprising parvovirus rep coding sequences and
XX PT parvovirus cap coding sequences, useful in producing higher stocks of
XX PT hybrid parvovirus vectors for delivering therapeutic nucleic acids to a
XX PT subject.
XX PS Disclosure; SEQ ID NO 25; 115pp; English.
XX PN
CC The present invention describes a polynucleotide (I), comprising
CC parvovirus rep coding sequences and parvovirus cap coding sequences. The
CC rep coding sequences encodes a DNA binding domain from a first
CC parvovirus, and a capsid interacting domain from a parvovirus different
CC from the first parvovirus. The cap coding sequence comprises sequences
CC from the different parvovirus. Also described: (1) a vector comprising
CC (1); (2) a cell comprising (1), or parvovirus rep coding sequences and
CC parvovirus cap coding sequences, where the rep coding sequences encode a
CC DNA binding domain from a first parvovirus and a capsid interacting
CC domain from a parvovirus different from the first parvovirus, the cap
CC coding sequences comprise sequences from the different parvovirus, and
CC the rep coding sequences are stably integrated into the genome of the
CC cell; and (3) producing a recombinant hybrid parvovirus particle or adeno
CC -associated virus (rAAV) particle. (I) can be used in vaccines, and in
CC gene therapy. The polynucleotide (I) can be used in producing higher
CC stocks of hybrid parvoviruses or parvovirus vectors, which may be used in
CC the delivery of nucleic acids having biological effect to treat or
CC ameliorate the symptoms associated with any disorder related to gene
CC expression. The polynucleotide may be used to produce a parvovirus vector
CC to express an immunogenic polypeptide in a subject, e.g. for vaccination.
CC The parvovirus vector may also be used to provide an antisense nucleic
CC acid to a cell in vitro or in vivo, or in diagnostic and screening
CC methods. The present sequence is used in the exemplification of the
CC present invention.
SQ Sequence 4683 BP; 1114 A; 1363 C; 1277 G; 929 T; 0 U; 0 Other;
Query Match 85.2%; Score 121.8; DB 12; Length 4683;
Best Local Similarity 98.4%; Pred. No. 1.3e-22;
Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 19 TTGCCCACTCCTCTCTGCGGCTCGCTCGCTCGGTGGGGCTTGGGACCAAGGTCCGC 78
DB 4683 TTGCCCACTCCTCTATGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCAAGGTCCGC 4624
QY 79 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 138
DB 4623 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 4564
QY 139 GGCAA 143
DB 4563 GGCAA 4559
RESULT 13
AAD00781/c
ID AAD00781 standard; DNA; 143 BP.
XX AC AAD00781;
XX DT 08-SEP-2000 (first entry)
XX DE Adeno-associated virus serotype 1 5' inverted terminal repeat.
XX KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
XX KW cap protein; recombinant viral vector; gene delivery; gene therapy;
XX KW vaccine; transgene; inverted terminal repeat; ITR; ss.
XX OS Adeno associated virus serotype 1.
XX PN
FH Key Location/Qualifiers
FT misc_binding 1..41
FT /tag= a
FT /bound_moiety= "binds to nucleotides 85..125"
FT stem_loop 42..62
FT /tag= b
FT stem_loop 64..84
FT /tag= c
FT misc_binding 85..125
FT /tag= d
FT /bound_moiety= "binds to nucleotides 41..1"
XX PN WO200028061-A2.
XX PD 18-MAY-2000.
XX PF 02-NOV-1999; 99WO-US025694.
XX PR 05-NOV-1998; 98US-0107114P.
XX PA (UYPE-) UNIV PENNSYLVANIA.
XX PI Wilson JM, Xiao W;
XX DR WPI; 2000-376571/32.
XX PT Novel adeno-associated virus serotype 1 polynucleotide useful for
XX PT preparation of medicament for delivery of a transgene to a host.
```

```
PS Example 2; Fig 2; 108pp; English.
XX
CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA
CC which is characterised by two inverted terminal repeats (ITR) and open
CC reading frames for rep and capsid (cap) proteins. The rep reading frame
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap
CC reading frame encodes three structural proteins, VP1, VP2 and VP3. The
CC AAV-1 sequence or its fragments particularly ITRs, rep and cap coding
CC regions, are useful in production of recombinant viral vectors for gene
CC delivery. These vectors can be used as gene therapy vectors, vaccine
CC vectors or antisense delivery vectors. The AAV-1 does not induce the
CC formation of neutralising antibodies specific to any serotype of AAV
CC hence is useful for transforming host cells, and in the preparation of a
CC medicament for the delivery of transgene to a host. The present sequence
CC is a 5' ITR of AAV-1 DNA which is useful in the production of recombinant
CC viral vector. The ITR forms a T-shaped hairpin structure
XX
SQ Sequence 143 BP; 24 A; 50 C; 48 G; 21 T; 0 U; 0 Other;
Query Match 83.2%; Score 119; DB 3; Length 143;
Best Local Similarity 89.5%; Pred. No. 5.9e-22;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
OY 1 TTACCCCTAGTGTGAGTTGCCACTCCCTCTCTCGCGGCTCGCTCGCTCGGTGGGGCC 60
DB 143 TTACCCCTAGTGTGAGTTGCCACTCCCTCTCTCGCGGCTCGCTCGCTCGGTGGGGCC 84
OY 61 TCGGACCAAGCTCCGACAGCGAGCTCTCTCTCGCGGCTCGCTCGCGCCACCGAGCGGCA 120
DB 83 GGCAGAGCAGAGCTCTGCGGCTCTGCGGACCTTTGGTCCGAGCGCCACCGAGCGGCA 24
OY 121 GCGCGCAGAGCGAGTGGGCAA 143
DB 23 GCGCGCAGAGCGAGTGGGCAA 1
RESULT 14
AAD00772/c
ID AAD00772 standard; DNA; 4718 BP.
XX
AC AAD00772;
XX
DT 08-SEP-2000 (first entry)
XX
DE Adeno-associated virus serotype 1 DNA.
XX
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
KW cap protein; recombinant viral vector; gene delivery; gene therapy;
KW vaccine; transgene; ss.
XX
OS Adeno associated virus serotype 1.
XX
FH Key Location/Qualifiers
FT repeat_unit 1..143
FT /tag= a
FT /label= 5' ITR.
FT /note= "Inverted terminal repeat which is capable of
FT forming T-shaped hairpin structure"
FT protein_bind 89..110
FT /tag= b
FT /bound_moiety= "Rep protein"
FT misc_feature 124..125
FT /tag= c
FT /note= "Terminal resolution site (TRS)"
FT misc_binding 219..226
FT /tag= d
FT /bound_moiety= "USP"
FT promoter 236..299
FT /tag= e
FT /label= P5_promoter
FT protein_bind 237..245
FT /tag= f
FT /bound_moiety= "YY1 factor"
FT TATA_signal 270..275
FT /tag= g
FT /label= P5_TATA-Box
FT misc_feature 299..306
FT /tag= h
FT /note= "YY1/P5 RNA"
FT CDS 335..2272
FT /product= "Rep 68"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT /note= "The coding region is interrupted by intron"
FT CDS 335..2206
FT /tag= i
FT /product= "Rep 78"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT TATA_signal 857..862
FT /tag= l
FT /label= P19_TATA_Box
FT misc_feature 882..883
FT /tag= m
FT /note= "P19 RNA"
FT CDS 1007..2272
FT /tag= o
FT /product= "Rep 40"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT /note= "The coding region is interrupted by intron"
FT CDS 1007..2206
FT /tag= n
FT /product= "Rep 52"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT TATA_signal 1836..1841
FT /tag= p
FT /label= P40_TATA-Box
FT misc_feature 1875..1876
FT /tag= q
FT /note= "P40 RNA"
FT intron 1924..2220
FT /tag= k
FT /note= "This region interrupts the coding sequence of Rep
FT 68 and Rep 40"
FT CDS 2223..4433
FT /tag= r
FT /product= "VP1 protein"
FT /note= "Capsid protein"
FT CDS 2634..4433
FT /tag= s
FT /product= "VP2 protein"
FT /note= "Capsid protein"
FT CDS 2829..4433
FT /tag= t
FT /product= "VP3 protein"
FT /note= "Capsid protein"
FT polyA_signal 4447..4452
FT /tag= u
FT repeat_unit 4576..4718
FT /tag= v
FT /label= 3' ITR
FT /note= "Inverted terminal repeat which is capable of
FT forming T-shaped hairpin structure"
XX
PN WO200028061-A2.
XX
PD 18-MAY-2000.
XX
PF 02-NOV-1999; 99WO-US025694.
XX
PR 05-NOV-1998; 98US-0107114P.
```

```
XX (UY28-) UNIV PENNSYLVANIA.
XX PA
XX PI Wilson JM, Xiao W;
XX XX
XX WPI; 2000-376571/32.
XX DR P-PSDB; AAV71161, AAV71164, AAV71165, AAV71166, AAV71167, AAV71168,
XX DR AAY71169.
XX XX
XX Novel adeno-associated virus serotype 1 polynucleotide useful for
XX PT preparation of medicament for delivery of a transgene to a host.
XX PS Claim 1; Fig 1; 108pp; English.
XX XX
XX The present sequence is an adeno-associated virus serotype 1 (AAV-1) DNA
XX CC characterised by two inverted terminal repeats (ITR) and open reading
XX CC frames for rep and capsid (cap) proteins. The rep reading frame encodes
XX CC four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading
XX CC frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1
XX CC sequence or its fragments particularly ITRs, rep and cap coding regions,
XX CC are useful in production of recombinant viral vectors for gene delivery.
XX CC These vectors can be used as gene therapy vectors, vaccine vectors or
XX CC antisense delivery vectors. The AAV-1 does not induce the formation of
XX CC neutralising antibodies specific to any serotype of AAV hence is useful
XX CC for transforming host cells, and in the preparation of a medicament for
XX CC the delivery of transgene to a host
XX XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
XX SQ
XX Query Match 83.2%; Score 119; DB 3; Length 4718;
XX Best Local Similarity 89.5%; Pred. No. 7.2e-22;
XX Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
XX QY 1 TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 60
XX Db 143 TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGCC 84
XX QY 61 TCGGACCAAGATCGCGAGAGCGGAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGA 120
XX Db 83 GGCAGAGCAGAGCTCTGCGCTCTGCGGACCTTTGGTCCGACGGCCGCCACCGAGCGAGCGA 24
XX QY 121 GCGCGCAGAGAGGAGTGGGCAA 143
XX Db 23 GCGCGCAGAGAGGAGTGGGCAA 1
XX RESULT 15
XX ID ADE76507/c
XX AC ADE76507 standard; DNA; 4718 BP.
XX AC ADE76507;
XX XX
XX 29-JAN-2004 (first entry)
XX XX
XX Adeno-associated virus (AAV) related DNA, SEQ ID No 6.
XX XX
XX adeno-associated virus; AAV; cytostatic; antipsoriatic; antirheumatic;
XX KW antiarthritic; neuroprotective; antidiabetic; antithyroid;
XX KW dermatological; antiinflammatory; gene therapy; vaccine;
XX KW hyperproliferative; cancer; psoriasis; autoimmune disease;
XX KW rheumatoid arthritis; multiple sclerosis; diabetes;
XX KW autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.
XX XX
XX Adeno-associated virus 1.
XX OS
XX BP1310571-A2.
XX PN
XX XX
XX 14-MAY-2003.
XX PD
XX XX
XX 13-NOV-2002; 2002EP-00257826.
XX PF
XX 13-NOV-2001; 2001US-0350607P.
XX PR
XX 17-DEC-2001; 2001US-0341117P.
XX PR
```

```
PR 01-MAY-2002; 2002US-0377066P.
PR 05-JUN-2002; 2002US-0386675P.
XX XX
XX (UY28-) UNIV PENNSYLVANIA.
XX PA
XX PI Gao G, Wilson JM, Alvira M;
XX XX
XX WPI; 2003-450984/43.
XX DR
XX XX
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.
XX PT preventing or treating hyperproliferative or autoimmune diseases,
XX PT comprises subjecting a sample having a DNA to amplification via
XX PT polymerase chain reaction.
XX PS Claim 14; SEQ ID NO 6; 419pp; English.
XX XX
XX The invention relates to a novel method for detecting adeno-associated
XX CC virus (AAV) sequences in a sample, which comprises subjecting a sample
XX CC containing a DNA to amplification via a polymerase chain reaction (PCR).
XX CC The AAV sequence have the following activities: cytostatic,
XX CC antipsoriatic, antirheumatic, antiarthritic, neuroprotective,
XX CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
XX CC sequence can be used in gene therapy or as part of a vaccine to treat
XX CC disorders. The method is useful in detecting and/or identifying AAV
XX CC sequences and isolating novel sequences that are identified. The
XX CC sequences may be used e.g. for preventing or treating hyperproliferative
XX CC conditions such as cancers and psoriasis, and other autoimmune diseases
XX CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
XX CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence
XX CC represents an AAV related DNA sequence of the invention.
XX XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
XX SQ
XX Query Match 83.2%; Score 119; DB 10; Length 4718;
XX Best Local Similarity 89.5%; Pred. No. 7.2e-22;
XX Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
XX QY 1 TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGCC 60
XX Db 143 TTACCCCTAGTGATGGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGCC 84
XX QY 61 TCGGACCAAGATCGCGAGAGCGGAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGA 120
XX Db 83 GGCAGAGCAGAGCTCTGCGCTCTGCGGACCTTTGGTCCGAGGCCGCCACCGAGCGAGCGA 24
XX QY 121 GCGCGCAGAGAGGAGTGGGCAA 143
XX Db 23 GCGCGCAGAGAGGAGTGGGCAA 1
XX Search completed: November 28, 2005, 19:55:37
XX Job time : 66.9124 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 12:51:27 ; Search time 522.566 Seconds  
(without alignments)  
15555.198 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_4576\_4718  
Perfect score: 143  
Sequence: 1 ttaccctcagtgatggagtt.....cgcagagaggagtgggcaa 143

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

ALIGNMENTS

GenEmbl: \*  
1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sta.\*  
11: gb\_sy.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_hcg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	100.0	4718	6	BD242766 Adeno-ss
2	143	100.0	4718	6	AR562498 Sequence
3	143	100.0	4718	6	AX753251 Sequence
4	143	100.0	4718	13	AF063497 Adeno-ss
5	135.6	94.8	4721	6	CQ972062 Sequence
6	135.6	94.8	4721	6	CS073592 Sequence
7	135.6	94.8	4721	6	AX753246 Sequence
8	135.6	94.8	4721	13	AF513851 Adeno-ss
9	121.8	85.2	4683	13	AF028704 Adeno-ss
10	120.2	84.1	4683	6	BD242775 Adeno-ss
11	120.2	84.1	4683	6	AR562507 Sequence
12	119	83.2	4718	6	BD242766 Adeno-ss
13	119	83.2	4718	6	AR562498 Sequence
14	119	83.2	4718	6	AX753251 Sequence
15	119	83.2	4718	13	AF063497 Adeno-ss
16	114.8	80.3	4683	13	AF028704 Adeno-ss
17	113.2	79.2	4683	6	BD242775 Adeno-ss
18	113.2	79.2	4683	6	AR562507 Sequence

19	113.2	79.2	4726	6	AX753253 Sequence
20	113.2	79.2	4726	13	AVU48704 Adeno-ss
c 21	111.6	78.0	4721	6	CQ972062 Sequence
c 22	111.6	78.0	4721	6	CS073592 Sequence
c 23	111.6	78.0	4721	6	AX753246 Sequence
c 24	111.6	78.0	4721	13	AF513851 Adeno-ss
25	109	76.2	4681	6	BD242774 Adeno-ss
26	109	76.2	4681	6	AR562506 Sequence
c 27	107.4	75.1	145	6	AR140333 Sequence
28	107.4	75.1	145	6	AX286293 Sequence
29	107.4	75.1	145	13	AA2L7R2 Adeno-ss
30	107.4	75.1	165	6	AR034135 Sequence
31	107.4	75.1	165	6	BD218219 Adeno-ss
32	107.4	75.1	165	6	II6806 Sequence 1
33	107.4	75.1	165	6	AR223306 Sequence
34	107.4	75.1	165	6	AR658389 Sequence
35	107.4	75.1	165	6	AX106702 Sequence
36	107.4	75.1	192	6	A46400 Sequence 3
c 37	107.4	75.1	198	6	AX418199 Sequence
38	107.4	75.1	207	6	AX703496 Sequence
39	107.4	75.1	272	6	BD218218 Adeno-ss
40	107.4	75.1	272	6	AR223305 Sequence
41	107.4	75.1	272	6	AR658388 Sequence
42	107.4	75.1	272	6	AX106701 Sequence
43	107.4	75.1	4675	6	BD094552 Method of
44	107.4	75.1	4675	6	AX135805 Sequence
45	107.4	75.1	4675	6	AX286292 Sequence

RESULT 1  
BD242766 4718 bp DNA linear PAT 17-JUL-2003  
LOCUS Adeno-associated virus serum type 1 nucleic acid sequence, vector  
DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector  
ACCESSION BD242766  
VERSION BD242766.1 GI:33052536  
KEYWORDS JP 2002529098-A/1.  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 4718)  
AUTHORS Wilson,J.M. and Xiao,W.  
TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector  
JOURNAL Patent: JP 2002529098-A 1 10-SEP-2002;  
THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA  
COMMENT OS AAV-1  
PN JP 2002529098-A/1  
PD 10-SEP-2002  
PF 02-NOV-1999 JP 2000581227  
PR 05-NOV-1998 US 60/107114  
PI JAMES M WILSON,WEIDONG XIAO  
PC C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC  
C12N1/21,  
PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC  
Adeno-associated virus serum type 1 nucleic acid sequence, CC  
vector and host  
CC cell containing the same  
FH Key Location/Qualifiers  
FT CDS (335) . . (2206)  
FT CDS (2223) . . (4430) .

FEATURES  
source

ORIGIN

Query Match 100.0%; Score 143; DB 6; Length 4718;  
Best Local Similarity 100.0%; Pred. NO. 1.8e-21;

```
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTATGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 60
DB 4576 TTACCCCTAGTATGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 4635
QY 61 TCGGACCAAAAGTTCGGCAGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGA 120
DB 4636 TCGGACCAAAAGTTCGGCAGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGGAGTGGGCAA 143
DB 4696 GCGCGCAGAGGAGTGGGCAA 4718

RESULT 2
AR562498
LOCUS AR562498 4718 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6759237.
ACCESSION AR562498
VERSION AR562498.1 GI:53976564
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4718)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors
and host cells containing same
JOURNAL Patent: US 6759237-A 1 06-JUL-2004;
The Trustees of The University of Pennsylvania; Philadelphia, PA
FEATURES
source
1. .4718
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 100.0%; Score 143; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 1.8e-21;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTATGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 60
DB 4576 TTACCCCTAGTATGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 4635
QY 61 TCGGACCAAAAGTTCGGCAGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGA 120
DB 4636 TCGGACCAAAAGTTCGGCAGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGGAGTGGGCAA 143
DB 4696 GCGCGCAGAGGAGTGGGCAA 4718

RESULT 4
AF063497
LOCUS AF063497 4718 bp DNA linear VRL 27-APR-1999
DEFINITION Adeno-associated virus 1, complete genome.
ACCESSION AF063497
VERSION AF063497.1 GI:4689096
KEYWORDS
SOURCE Adeno-associated virus 1
ORGANISM Adeno-associated virus 1
REFERENCE 1 (bases 1 to 4718)
AUTHORS Xiao,W., Chirmule,N., Berta,S.C., McCullough,B., Gao,G. and Wilson,J.M.
TITLE Gene therapy vectors based on adeno-associated virus type 1
JOURNAL J. Virol. 73 (5), 3994-4003 (1999)
PUBMED 10196295
REFERENCE 2 (bases 1 to 4718)
AUTHORS Xiao,W. and Wilson,J.M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601 Spruce Street, Philadelphia, PA 19104, USA
FEATURES
source
1. .4718
/organism="Adeno-associated virus 1"
/mol_type="genomic DNA"
/db_xref="taxon:85106"
/note="AAV1"
335. .2206
/codon_start=1
CDS
/product="nonstructural protein"
/protein_id="AAD27758.1"
/db_xref="GI:4689098"
/translaton="MPGYEIVIKVPSDLDEHLPGISDFVSWVAEKWEWLPDSDMD
LNLIQAPLVAEKLQDFLVQMRVSKAPALFEVQPKGESYFHLHLVETTVKVS
MYLGRFLSQIRDLKVLQTYIRGIEPLFPLNFAVTKRNGAGGNGKVVDSYIPNLYLPK
TQPELQWQNMMEYISACLNLAERKLVAQLTHVSTQSQNKENLNPNSDAPVRS
TKSADPYLVGPAPPADIKTNRITRIELNGYEPAYAGSVFLGWAQKRFKRNITWLF
PATTKNTIAEALAHAVPFYGCVMNTNFPNCDVKNVIMWEEGKMTAKVVEGAKA
ILGSKVRVDOKSSSAOI DPTPTVITNTNMCVVIDGNSTTFEHOQPLQDRMFKPEL
TRLEHDFGKVTQKVEFFRMAQDHVTEVAHEFYVRKGANKRPAPODADKSEKRA
RCSVADPSTSDAEGAPVDFADRYQNKSRHAGMLQMLFPCKTCERMNQFNLCPTHGT
RDCSCFFGVGSSEQPVVRKRTYRKLCAIHLHGRAPEIACACDLVNVLDLDCVSEQ"
2223. .4433
/codon_start=1
CDS
/product="capsid protein"
/protein_id="AAD27757.1"
/db_xref="GI:4689097"
/translaton="MAADGYLPDWLEDNLSGIREWBDLKPQAPKPKANQKQDDGRG
LVLPGYKYLGPNGLDKGEVNAADAAALHDKAYDQQLKAGDNPYLRNHNADBSFQ
RIQEDTSFGNLRGAVFQAKKRVLEPLGLVBEGAKTAPGKKRPVQSPQEPDSSSGIQ
KTGQQQPAKKRLNFGQTGDSSEVPDPQPLGEPATPAAVGPTTMSGGGAPMADNNEGA
```

```
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTATGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 60
DB 4576 TTACCCCTAGTATGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 4635
QY 61 TCGGACCAAAAGTTCGGCAGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGA 120
DB 4636 TCGGACCAAAAGTTCGGCAGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGGAGTGGGCAA 143
DB 4696 GCGCGCAGAGGAGTGGGCAA 4718

RESULT 2
AR562498
LOCUS AR562498 4718 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6759237.
ACCESSION AR562498
VERSION AR562498.1 GI:53976564
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4718)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors
and host cells containing same
JOURNAL Patent: US 6759237-A 1 06-JUL-2004;
The Trustees of The University of Pennsylvania; Philadelphia, PA
FEATURES
source
1. .4718
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Query Match 100.0%; Score 143; DB 6; Length 4718;
Best Local Similarity 100.0%; Pred. No. 1.8e-21;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTATGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 60
DB 4576 TTACCCCTAGTATGAGTTGCCACTCCCTCTCTGCGGCTCGCTCGCTCGGTGGGCC 4635
QY 61 TCGGACCAAAAGTTCGGCAGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGA 120
DB 4636 TCGGACCAAAAGTTCGGCAGACGGCAGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGA 4695
QY 121 GCGCGCAGAGGAGTGGGCAA 143
DB 4696 GCGCGCAGAGGAGTGGGCAA 4718

RESULT 3
AX753251
LOCUS AX753251 4718 bp DNA linear PAT 23-JUN-2003
DEFINITION Sequence 6 from Patent EP1310571.
ACCESSION AX753251
VERSION AX753251.1 GI:32166108
KEYWORDS
SOURCE Adeno-associated virus 1
ORGANISM Adeno-associated virus 1
REFERENCE 1
AUTHORS Gao,G., Wilson,J.M. and Alvira,M.
TITLE A method of detecting and/or identifying adeno-associated virus
(AAV) sequences and isolating novel sequences identified thereby
JOURNAL Patent: EP 1310571-A 6 14-MAY-2003;
The Trustees of The University of Pennsylvania (US)
FEATURES
source
1. .4718
/organism="Adeno-associated virus 1"
/mol_type="genomic DNA"
/db_xref="taxon:85106"
/note="AAV1"
335. .2206
/codon_start=1
CDS
/product="nonstructural protein"
/protein_id="AAD27758.1"
/db_xref="GI:4689098"
/translaton="MPGYEIVIKVPSDLDEHLPGISDFVSWVAEKWEWLPDSDMD
LNLIQAPLVAEKLQDFLVQMRVSKAPALFEVQPKGESYFHLHLVETTVKVS
MYLGRFLSQIRDLKVLQTYIRGIEPLFPLNFAVTKRNGAGGNGKVVDSYIPNLYLPK
TQPELQWQNMMEYISACLNLAERKLVAQLTHVSTQSQNKENLNPNSDAPVRS
TKSADPYLVGPAPPADIKTNRITRIELNGYEPAYAGSVFLGWAQKRFKRNITWLF
PATTKNTIAEALAHAVPFYGCVMNTNFPNCDVKNVIMWEEGKMTAKVVEGAKA
ILGSKVRVDOKSSSAOI DPTPTVITNTNMCVVIDGNSTTFEHOQPLQDRMFKPEL
TRLEHDFGKVTQKVEFFRMAQDHVTEVAHEFYVRKGANKRPAPODADKSEKRA
RCSVADPSTSDAEGAPVDFADRYQNKSRHAGMLQMLFPCKTCERMNQFNLCPTHGT
RDCSCFFGVGSSEQPVVRKRTYRKLCAIHLHGRAPEIACACDLVNVLDLDCVSEQ"
2223. .4433
/codon_start=1
CDS
/product="capsid protein"
/protein_id="AAD27757.1"
/db_xref="GI:4689097"
/translaton="MAADGYLPDWLEDNLSGIREWBDLKPQAPKPKANQKQDDGRG
LVLPGYKYLGPNGLDKGEVNAADAAALHDKAYDQQLKAGDNPYLRNHNADBSFQ
RIQEDTSFGNLRGAVFQAKKRVLEPLGLVBEGAKTAPGKKRPVQSPQEPDSSSGIQ
KTGQQQPAKKRLNFGQTGDSSEVPDPQPLGEPATPAAVGPTTMSGGGAPMADNNEGA
```

DGVGNASGWHCDSTWLGDRVITSTRTWALPTYNHLYKQISSASTGASNDNHVFGY  
STPMGYFDNRHCHFSRDRWRLNNMNGFRKRLNFKLNIQKVEVTTNDGVTTIA  
NNLTSTVQVFSSEYQLPYVLGSAHQGLPPFADVFMIPQYGYLTLANGSOAVGRSS  
FYCLEVFPQMLRTNFTFSVTFEVPPHSHYSQSLDRMLNPLIDLYLNRTO  
NOSGAONKDLFPSGSPAGMSVQPKNMLPGPCYRQORVSKTKTNNNSNFTWTGASK  
YNLNGRESLINFCTAMASHKDDDEKFFPMSGVNMIKESAGASNTALDNVMTDSEEI  
KATNPVATERFTVAVNFOSSTDTATGDVHWAGALPGMVWQORDYILOGLWAKIPH  
TDGHPHPSLMGGFGLKNPQILIKNTVPANPPAFBSATKFAFPIQYSGQVSVE  
IEWELQKENSKEWNEVQYTSYAKSANYDFTVDNNGLYTEPRPIGTRYLTRPL"

ORIGIN

Query Match 100.0%; Score 143; DB 13; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 1.8e-21;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTCGCGCTCGCTCGCTCGGTGGGGCC 60  
|||||  
Db 4576 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTCGCGCTCGCTCGCTCGGTGGGGCC 4635  
|||||

QY 61 TCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTCGCGCCCCACCGAGCGAGCGA 120  
|||||  
Db 4636 TCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTCGCGCCCCACCGAGCGAGCGA 4695  
|||||

QY 121 GCGCGCAGAGGGAGTGGGCAA 143  
|||||  
Db 4696 GCGCGCAGAGGGAGTGGGCAA 4718  
|||||

RESULT 5  
CQ972062  
LOCUS 4721 bp DNA linear PAT 05-JAN-2005  
DEFINITION Sequence 7 from Patent WO2004108922.  
ACCESSION CQ972062  
VERSION CQ972062.1 GI:57163375  
KEYWORDS  
SOURCE Adeno-associated virus  
ORGANISM Adeno-associated virus  
1 Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.

REFERENCE  
AUTHORS Rader, D.J. and Wilson, J.M.  
TITLE Methods and compositions for lowering total cholesterol levels and treatment of heart disease  
JOURNAL Patent: WO 2004108922-A 7 16-DEC-2004;  
The Trustees of The University of Pennsylvania (US)  
FEATURES  
source Location/Qualifiers  
1. .4721  
/organism="Adeno-associated virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:272636"

ORIGIN

Query Match 94.8%; Score 135.6; DB 6; Length 4721;  
Best Local Similarity 97.2%; Pred. No. 7.2e-20;  
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCCCT 61  
|||||  
Db 4580 TACCCCTAGTGATGAGTTGCCACTCCCTCTATGCGCGCTCGCTCGCTCGGTGGGGCCCT 4639  
|||||

QY 62 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGAG 121  
|||||  
Db 4640 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGAG 4699  
|||||

QY 122 GCGCGCAGAGGGAGTGGGCAA 143  
|||||  
Db 4700 GCGCGCAGAGGGAGTGGGCAA 4721  
|||||

RESULT 6  
CQ973592  
LOCUS 4721 bp DNA linear PAT 05-MAY-2005  
DEFINITION Sequence 180 from Patent WO200503321.  
ACCESSION CQ973592

CS073592.1 GI:63090489  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1  
AUTHORS Wilson, J.M., Gao, G., Alvira, M.R. and Vandenbergh, L.H.  
TITLE Adeno-associated virus (aav) clades, sequences, vectors containing same, and uses therefor  
JOURNAL Patent: WO 200503321-A 180 14-APR-2005;  
The Trustees of the University of Pennsylvania (US)  
FEATURES  
source Location/Qualifiers  
1. .4721  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="adeno-associated virus serotype 7"

ORIGIN

Query Match 94.8%; Score 135.6; DB 6; Length 4721;  
Best Local Similarity 97.2%; Pred. No. 7.2e-20;  
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCCCT 61  
|||||  
Db 4580 TACCCCTAGTGATGAGTTGCCACTCCCTCTATGCGCGCTCGCTCGCTCGGTGGGGCCCT 4639  
|||||

QY 62 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGAG 121  
|||||  
Db 4640 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGAG 4699  
|||||

QY 122 GCGCGCAGAGGGAGTGGGCAA 143  
|||||  
Db 4700 GCGCGCAGAGGGAGTGGGCAA 4721  
|||||

RESULT 7  
AX753246  
LOCUS 4721 bp DNA linear PAT 23-JUN-2003  
DEFINITION Sequence 1 from Patent EP1310571.  
ACCESSION AX753246  
VERSION AX753246.1 GI:32166105  
KEYWORDS  
SOURCE Adeno-associated virus 7  
ORGANISM Adeno-associated virus 7  
1 Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.

REFERENCE  
AUTHORS Gao, G., Wilson, J.M. and Alvira, M.  
TITLE A method of detecting and/or identifying adeno-associated virus (AAV) sequences and isolating novel sequences identified thereby  
JOURNAL Patent: EP 1310571-A 1 14-MAY-2003;  
The Trustees of The University of Pennsylvania (US)  
FEATURES  
source Location/Qualifiers  
1. .4721  
/organism="Adeno-associated virus 7"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:202812"

ORIGIN

Query Match 94.8%; Score 135.6; DB 6; Length 4721;  
Best Local Similarity 97.2%; Pred. No. 7.2e-20;  
Matches 138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCCCT 61  
|||||  
Db 4580 TACCCCTAGTGATGAGTTGCCACTCCCTCTATGCGCGCTCGCTCGCTCGGTGGGGCCCT 4639  
|||||

QY 62 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGAG 121  
|||||  
Db 4640 GCGGACCAAAAGTCCGACAGCGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGAG 4699  
|||||

QY 122 GCGCGCAGAGGGAGTGGGCAA 143  
|||||  
Db 4700 GCGCGCAGAGGGAGTGGGCAA 4721  
|||||





FYCLRYFSPQMLRTGNFTSYTFEDVPHSSYAHQSQSLRLMNLPLIDQYLYLNRTO
NQGSAQNKLDFSRGSPAGMSVQPKMLPGCYRQVRVSKTKTNNNSNFTWTGASK
YNLNGRESTINPCTAMASHKDDKFFPMGVMYI FGESAGASNTALDNVMTDEEII
KATNPVATERFCTAVNLQSSSTPATGTVHVMGALPGVMODRDVYLGQPTWAKIPH
TUGHFSPKMGGLGKHPPOILLKNTFPVPAEPAESATFASFTQYSTQVSQVSE
ISEWLOKSKRWNPVEVQYTSYAKSANVDFTVDNNGLYTEPRPIGTRILIRPL"

ORIGIN

Query Match 85.2%; Score 121.8; DB 13; Length 4683;  
 Best Local Similarity 98.4%; Pred. No. 7.1e-17;  
 Matches 123; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 19 TTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTGCGACCAAGGTCCGC 78  
 DB 4683 TTGCCACTCCCTCTATGCGCGCTCGCTCGCTCGGTGGGGCTGCGACCAAGGTCCGC 4624  
 QY 79 AGACGGCAGAGCTCTGCTTGGCGGCCCCACCGAGCGAGCGCGCAGAGGAGTG 138  
 DB 4623 AGACGGCAGAGCTCTGCTTGGCGGCCCCACCGAGCGAGCGCGCATAGAGGAGTG 4564  
 QY 139 GGCAA 143  
 DB 4563 GGCAA 4559

RESULT 10  
BD242775/c

LOCUS BD242775 4683 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector  
 and host cell containing the same.  
 ACCESSION BD242775  
 VERSION JP 2002529098-A/10.  
 KEYWORDS unidentified  
 SOURCE unidentified  
 ORGANISM unclassified.  
 REFERENCE 1 (bases 1 to 4683)  
 AUTHORS Wilson,J.M. and Xiao,W.  
 TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector  
 and host cell containing the same  
 JOURNAL Patent: JP 2002529098-A 10 10-SEP-2002;  
 THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA  
 COMMENT OS aav-6  
 PN JP 2002529098-A/10  
 PD 10-SEP-2002  
 PF 02-NOV-1999 JP 2000581227  
 PR 05-NOV-1998 US 60/107114  
 PI JAMES M WILSON,WEIDONG XIAO  
 PC C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC  
 C12N1/21,  
 PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC  
 Adeno-associated virus serum type 1 nucleic acid sequence, CC  
 vector and host  
 CC cell containing the same  
 FH key Location/Qualifiers  
 FT source 1. 4683  
 FT /organism='aav-6'.  
 FEATURES  
 source  
 1. 4683  
 /organism='unidentified'  
 /mol\_type='genomic DNA'  
 /db\_xref='taxon:32644'

ORIGIN

Query Match 84.1%; Score 120.2; DB 6; Length 4683;  
 Best Local Similarity 97.6%; Pred. No. 1.6e-16;  
 Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 19 TTGCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCTGCGACCAAGGTCCGC 78  
 DB 4683 TTGCCACTCCCTCTATGCGCGCTCGCTCGGTGGGGCTGCGACCAAGGTCCGC 4624  
 QY 79 AGACGGCAGAGCTCTGCTTGGCGGCCCCACCGAGCGAGCGCGCAGAGGAGTG 138

Db 4623 AGACGGCAGAGCTCTGCTTGGCGGCCCCACCGAGCGAGCGCGCATAGAGGAGTG 4564  
 QY 139 GGCAA 143  
 DB 4563 GGCAA 4559

RESULT 11  
 AR562507/c  
 LOCUS AR562507 4683 bp DNA linear PAT 08-OCT-2004  
 DEFINITION Sequence 19 from patent US 6759237.  
 ACCESSION AR562507  
 VERSION AR562507.1 GI:53976573  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 4683)  
 AUTHORS Wilson,J.M. and Xiao,W.  
 TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors  
 and host cells containing same  
 JOURNAL Patent: US 6759237-A 19 06-JUL-2004;  
 The Trustees of the University of Pennsylvania; Philadelphia, PA  
 FEATURES  
 Location/Qualifiers  
 source  
 1. 4683  
 /organism='unknown'  
 /mol\_type='genomic DNA'

ORIGIN

Query Match 84.1%; Score 120.2; DB 6; Length 4683;  
 Best Local Similarity 97.6%; Pred. No. 1.6e-16;  
 Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 19 TTGCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCTGCGACCAAGGTCCGC 78  
 DB 4683 TTGCCACTCCCTCTATGCGCGCTCGCTCGGTGGGGCTGCGACCAAGGTCCGC 4624  
 QY 79 AGACGGCAGAGCTCTGCTTGGCGGCCCCACCGAGCGAGCGCGCAGAGGAGTG 138  
 DB 4623 AGACGGCAGAGCTCTGCTTGGCGGCCCCACCGAGCGAGCGCGCATAGAGGAGTG 4564  
 QY 139 GGCAA 143  
 DB 4563 GGCAA 4559

RESULT 12  
BD242766/c

LOCUS BD242766 4718 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector  
 and host cell containing the same.  
 ACCESSION BD242766  
 VERSION BD242766.1 GI:33052536  
 KEYWORDS JP 2002529098-A/1.  
 SOURCE unidentified  
 ORGANISM unidentified  
 REFERENCE 1 (bases 1 to 4718)  
 AUTHORS Wilson,J.M. and Xiao,W.  
 TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector  
 and host cell containing the same  
 JOURNAL Patent: JP 2002529098-A 1 10-SEP-2002;  
 THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA  
 COMMENT OS AAV-1  
 PN JP 2002529098-A/1  
 PD 10-SEP-2002  
 PF 02-NOV-1999 JP 2000581227  
 PR 05-NOV-1998 US 60/107114  
 PI JAMES M WILSON,WEIDONG XIAO  
 PC C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC  
 C12N1/21,  
 PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC  
 Adeno-associated virus serum type 1 nucleic acid sequence, CC  
 vector and host  
 CC cell containing the same  
 FH key Location/Qualifiers  
 FT source 1. 4683  
 FT /organism='aav-6'.  
 FEATURES  
 source  
 1. 4683  
 /organism='unidentified'  
 /mol\_type='genomic DNA'  
 /db\_xref='taxon:32644'

ORIGIN

Query Match 84.1%; Score 120.2; DB 6; Length 4683;  
 Best Local Similarity 97.6%; Pred. No. 1.6e-16;  
 Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 19 TTGCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCTGCGACCAAGGTCCGC 78  
 DB 4683 TTGCCACTCCCTCTATGCGCGCTCGCTCGGTGGGGCTGCGACCAAGGTCCGC 4624  
 QY 79 AGACGGCAGAGCTCTGCTTGGCGGCCCCACCGAGCGAGCGCGCAGAGGAGTG 138

```
Adeno-associated virus serum type 1 nucleic acid sequence, CC
vector and host
CC cell containing the same
FH Key Location/Qualifiers
FT CDS (335)..(2206)
FT CDS (2223)..(4430).
FEATURES
    source
        1..4718
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
ORIGIN
    Query Match      83.2%; Score 119; DB 6; Length 4718;
    Best Local Similarity 89.5%; Pred. No. 2.9e-16;
    Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 60
DB 143 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 84
QY 61 TCGGGACCAAGGTCGCGAGCGGAGAGCTCTGCTCTGCCCGGCCACCGAGCGAGCGA 120
DB 83 GCGAGAGCAGAGCTCTGCGCTCTGCGGACCTTTGGTTCGCGAGCGGCCACCGAGCGA 24
QY 121 GCGCGCAGAGAGGAGTGGGCAA 143
DB 23 GCGCGCAGAGAGGAGTGGGCAA 1
RESULT 13
AR562498/c
LOCUS AR562498 4718 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 1 from patent US 6759237.
ACCESSION AR562498
VERSION AR562498.1 GI:53976564
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4718)
AUTHORS Wilson,J.M. and Xiao,W.
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors
and host cells containing same
JOURNAL Patent: US 6759237-A 1 06-JUL-2004;
The Trustees of the University of Pennsylvania; Philadelphia, PA
FEATURES
    source
        1..4718
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
    Query Match      83.2%; Score 119; DB 6; Length 4718;
    Best Local Similarity 89.5%; Pred. No. 2.9e-16;
    Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 60
DB 143 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 84
QY 61 TCGGGACCAAGGTCGCGAGCGGAGAGCTCTGCTCTGCCCGGCCACCGAGCGAGCGA 120
DB 83 GCGAGAGCAGAGCTCTGCGCTCTGCGGACCTTTGGTTCGCGAGCGGCCACCGAGCGA 24
QY 121 GCGCGCAGAGAGGAGTGGGCAA 143
DB 23 GCGCGCAGAGAGGAGTGGGCAA 1
RESULT 14
AX753251/c
LOCUS AX753251 4718 bp DNA linear PAT 23-JUN-2003
DEFINITION Sequence 6 from Patent EP1310571.
```

```
ACCESSION AX753251
VERSION AX753251.1 GI:32166108
KEYWORDS Adeno-associated virus 1
SOURCE Adeno-associated virus 1
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1
AUTHORS Gao,G., Wilson,J.M. and Alvira,M.
TITLE A method of detecting and/or identifying adeno-associated virus
(AVV) sequences and isolating novel sequences identified thereby
JOURNAL Patent: EP 1310571-A 6 14-MAY-2003;
The Trustees of The University of Pennsylvania (US)
FEATURES
    Location/Qualifiers
        source
            1..4718
                /organism="Adeno-associated virus 1"
                /mol_type="unassigned DNA"
                /db_xref="taxon:85106"
ORIGIN
    Query Match      83.2%; Score 119; DB 6; Length 4718;
    Best Local Similarity 89.5%; Pred. No. 2.9e-16;
    Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 60
DB 143 TTACCCCTAGTGATGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGCC 84
QY 61 TCGGGACCAAGGTCGCGAGCGGAGAGCTCTGCTCTGCCCGGCCACCGAGCGAGCGA 120
DB 83 GCGAGAGCAGAGCTCTGCGCTCTGCGGACCTTTGGTTCGCGAGCGGCCACCGAGCGA 24
QY 121 GCGCGCAGAGAGGAGTGGGCAA 143
DB 23 GCGCGCAGAGAGGAGTGGGCAA 1
RESULT 15
AF063497/c
LOCUS AF063497 4718 bp DNA linear VRL 27-APR-1999
DEFINITION Adeno-associated virus 1, complete genome.
ACCESSION AF063497
VERSION AF063497.1 GI:4689096
KEYWORDS Adeno-associated virus 1
SOURCE Adeno-associated virus 1
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1 (bases 1 to 4718)
AUTHORS Xiao,W., Chirmule,N., Berta,S.C., McCullough,B., Gao,G. and
Wilson,J.M.
TITLE Gene therapy vectors based on adeno-associated virus type 1
JOURNAL J. Virol. 73 (5), 3994-4003 (1999)
PUBMED 10196295
REFERENCE 2 (bases 1 to 4718)
AUTHORS Xiao,W. and Wilson,J.M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601
Spruce Street, Philadelphia, PA 19104, USA
FEATURES
    Location/Qualifiers
        1..4718
            /organism="Adeno-associated virus 1"
            /mol_type="genomic DNA"
            /db_xref="taxon:85106"
            /note="AAV1"
            335..2206
                /codon_start=1
                /product="nonstructural protein"
                /protein_id="AAD2758.1"
                /db_xref="GI:4689098"
                /translation="MPGFYFIVIKVPSLDLHLPDISFVSFVVAEKSWELPPDSMD
                LNLIEQALPTVAELKLFQVWRVSKAPEALFFVQFEKGESYFHLHLIVETTVKKS
                MYLGRFLSQIRDKLVQTIYRGIEPTLPNFAVTKRNGAGGNGKVVDCIYINLLPK
                TOPELQAWNTMMEYISACLNAERKRLVAHLTHVSQTQEQNKENLPNSDAPVIRS
                KTSARYMELVGLVDRGITSEKQWIQEDQASYISFNAAASNSRSQIKALDINAGKIMAL"
CDS
```

TKSAPDYLVGAPPADIKTNRIYRIELENGYEPAYAGSVFLGWAQKRFGRNTIWLFG  
PATTKNTIAEAIAHAFYPCVNTNENFPFNDVCVRKVIWHEEGMTAKYVESAKA  
ILGSKVRVDQKCKSAQIDPTPIVITSNTNCAVIDGNSITTFHQOPLQDRMFKFEL  
TRLEHDFGKVKQEFERWAQDHVTEVAHEFYVRKGGANKRPAPDDADKSEPKRA  
CPSLADPSTDAEGAPVDPADRYQNKCSRHAGLQWLFCKTCERNQNFN1CFTHGT  
RGCSECFPGVSESQPVVRKTRKLCIAIHLLGRAPETIACSACDLVVDLDDCVSEQ"  
2223. .4433  
/codon\_start=1  
/product="capsid protein"  
/protein\_id="AAD27757.1"  
/db\_xref="GI:4689097"  
/translations="MAADGYLDPWLEDNLSEGREWDLKPGAPKPKANQKQDDGRG  
LVLPGYKILGPFNGKGEFVPAADAAALEHDKAYDQQLKAGDNPIRYNHADAEPQE  
RLQEDTSFGNLRGAVFQAKRVLEPLGVEGAKTAPGKRPEQSPQEPDSSSGIG  
KTGQQAQKRLNFGQTSVDPQPLGEPFATPAAGVPTTMASSGGA PMADNNHFGY  
DGVNAGSNWCHDSTWLGDRVITSTRWALPTYNHLYKQISSASTGASNDNHFGY  
STPMGYFDNRFHCHESPRDWORLNNNGFRPKRLNFKLNIQVKEVTNDGVTTIA  
NMLTSTQVFSDEYQLPYVLGSAHQCCLPPPADVFMIPQYGYLTLNNGSQAVERSS  
FYCLEYFQSOMLRTGNFTSYTFEEVFPFSSYAHQSILDRMLNPLIDQLYLYLNRTQ  
NOSGSAQNKDLIFSRGSPAGMSVQPKWLPFGCYRQORVSKTKTDNNNSNFTWTGASK  
YNLNGRESIINPGTAMASHKDDKDFPMSGVMI FGKESAGASNTALDNVMTDEEEI  
KATNPVATERFGTVAVNFOSSTDPATGVDHAMGLPMWODRDVYLOQPIWAKIPH  
TDGHFHPSLMCGFGLKNPPQIILIKNTVPANPPAEFSATKPFASFILOYSTGOVSVE  
IEWELQKNSKRWNPVQYTSNTAKSANVDFTVDNNGLYTEPRPIGTRILTRPL"

CDS

ORIGIN

Query Match 83.2%; Score 119; DB 13; Length 4718;  
Best Local Similarity 89.5%; Pred. No. 2.9e-16;  
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
QY 1 TTACCCCTAGTGTGAGTTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCC 60  
Db 143 TTAGCCCTAGTGTGAGTTGCCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCC 84  
QY 61 TCGGACCAAGTCCGACACGCGAGAGCTCTGCTCTGCGGCGCCACCGAGCGAGCGA 120  
Db 83 GGCAGAGCAGAGCTCTGCTCTGCGGACCTTTGGTCCGAGGCCCCACCGAGCGAGCGA 24  
QY 121 GCGCGCAGAGAGGAGTGGGCAA 143  
Db 23 GCGCGCAGAGAGGAGTGGGCAA 1

Search completed: November 29, 2005, 00:05:26  
Job time : 523.566 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 18:24:29 ; Search time 454.049 Seconds  
(without alignments)  
14735.325 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_4576\_4718

Perfect score: 143

Sequence: 1 ttaccctagtgatggagt.....cgacagagggagtgggcaa 143

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_hcc:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_est7:\*  
9: gb\_gse1:\*  
10: gb\_gse2:\*  
11: gb\_gse3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37.8	26.4	763	2	BG490707
C 2	37.8	26.4	1021	5	B0838990
C 3	37.4	26.2	644	2	BB632576
C 4	36.2	25.3	604	3	BI531460
C 5	36.2	25.3	932	10	AL066742
C 6	36.2	25.3	1021	5	B0838990
C 7	36.2	25.2	318	7	CV352119
C 8	35.8	25.0	1074	10	CL507458
C 9	35.6	24.9	578	2	BB651469
C 10	35.6	24.9	793	3	BI905802
C 11	35.6	24.9	1115	3	BM906771
C 12	35.6	24.8	1123	10	AG080476
C 13	35.2	24.6	2024	10	AG074763
C 14	35	24.5	724	10	CL751558
C 15	35	24.5	844	10	CN30052P
C 16	34.8	24.3	676	9	BZ414423
C 17	34.6	24.2	617	7	CN038418
C 18	34.6	24.2	637	7	CN045740
C 19	34.6	24.2	695	10	CZ814168
C 20	34.6	24.2	860	2	BF578052
C 21	34.6	24.2	926	8	CV775269
C 22	34.6	24.2	2088	10	CL507788

C 23	34.4	24.1	671	10	CH914921
C 24	34.4	24.1	890	10	AG127787
C 25	34.4	24.1	1136	3	BM416106
C 26	34.2	23.9	364	5	BY014749
C 27	34.2	23.9	468	8	CX358645
C 28	34.2	23.9	914	3	BI762701
C 29	34.2	23.9	1009	10	CNS010EW
C 30	34	23.8	343	5	BY181775
C 31	34	23.8	450	5	BY250974
C 32	34	23.8	623	2	BB613613
C 33	34	23.8	631	5	BY723537
C 34	34	23.8	639	2	BB651977
C 35	34	23.8	643	2	BB652152
C 36	34	23.8	657	2	BB651054
C 37	34	23.8	661	2	BB621081
C 38	34	23.8	678	5	BY726305
C 39	34	23.8	795	9	AQ331299
C 40	34	23.8	838	8	CV780024
C 41	34	23.8	1134	6	CA459340
C 42	34	23.8	1364	5	B0543298
C 43	34	23.8	1767	4	AK082810
C 44	34	23.8	2268	4	AK082478
C 45	34	23.8	2708	4	AK038836

ALIGNMENTS

RESULT 1  
BG490707/c  
LOCUS  
DEFINITION  
602520211f1 NIH\_MGC\_18 Homo sapiens cDNA clone IMAGE:4638699 5',  
mRNA sequence.  
ACCESSION  
BG490707  
VERSION  
BG490707.1  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
REFERENCE  
1 (bases 1 to 763)  
NIH-MGC http://mgc.mci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DPF/Gazdar  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCM1401 row: e column: 04  
High quality sequence stop: 182.

FEATURES  
Location/Qualifiers  
1..763  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4638699"  
/tissue\_type="large cell carcinoma"  
/lab\_host="PH108 (phage-resistant)"  
/clone\_lib="NIH\_MGC\_18"  
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 26.4%; Score 37.8; DB 2; Length 763;  
 Best Local Similarity 58.4%; Pred. No. 13;  
 Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 26 CTTCTCTCTGCGGCTCGCTCGTGGGGCTCGGGACCAAGGTCGCGACGCG 85  
 DB 261 CCGTGGCTGCGGGGCGCTCGCTCTGCTGCTGCGGCTCGCTTCGCGC 202  
 QY 86 AGAGCTCTGCTGCGGCGCCACCGAGCGAGCGCGCAGAGGAGTG 138  
 DB 201 AGCGCTACGCTCGCGGCGCTGACGCGGAGGCGGCTGTCGCGCGTG 149

## RESULT 2

LOCUS BU838990/c  
 DEFINITION AGENCOURT\_8209898 NIH\_MGC\_112 Homo sapiens cDNA clone IMAGE:6258289  
 5' mRNA sequence.

ACCESSION BU838990  
 VERSION BU838990.1 GI:24023385  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominiidae; Homo.

REFERENCE 1 (bases 1 to 1021)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: DCTP/DRP

cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LLCM2413 row: p column: 02  
 High quality sequence start: 36  
 High quality sequence stop: 231.

## FEATURES

Location/Qualifiers  
 1..1021  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6258289"  
 /tissue\_type="melanotic melanoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 112"  
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GCACAGAG(G). Library constructed by Ling Hong in the  
 Laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a  
 NIH\_MGC Library."

## ORIGIN

Query Match 26.4%; Score 37.8; DB 5; Length 1021;  
 Best Local Similarity 58.4%; Pred. No. 13;  
 Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 28 CCTCTCTGCGGCTCGCTCGTGGGGCTCGGACCAAGTCCGCGACGCGAG 87  
 DB 903 CCGGCGCGCGGCTAGCGGGGCGGGGCGGCGCTCCGCGCGCGTCCGGGGGCGCG 844  
 QY 88 AGCTCTGCTCTGCGGCGCCACCGAGCGAGCGAGCGCGCAGAGGAGTG 140

## Db

843 CGACGGGGGAAGCGCCCGCGGCGCTCGCGGGGTGGTGGGGG 791

RESULT 3  
 LOCUS BB632576

DEFINITION BB632576 RIKEN full-length enriched, adult male hypothalamus Mus  
 musculus cDNA clone A230067E15 5', mRNA sequence.

ACCESSION BB632576  
 VERSION BB632576.1 GI:16469164  
 KEYWORDS EST.

SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Muridea; Muridae; Murinae; Mus.

## REFERENCE

AUTHORS 1 (bases 1 to 644)  
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,  
 Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,  
 Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,  
 Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,  
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
 Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
 Unpublished (2001)

TITLE  
 JOURNAL  
 COMMENT

Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suihro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-3222  
 Fax: 81-45-503-9216

Email: [genome-resesec.riken.jp](mailto:genome-resesec.riken.jp), URL: <http://genome.gsc.riken.jp/>  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
 Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.  
 and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
 Sugahara, Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I.,  
 Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
 Hayashizaki, Y.  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
 further details.  
 e mouse tissues.

## FEATURES

Location/Qualifiers  
 1..644  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="A230067E15"  
 /sex="male"  
 /tissue\_type="hypothalamus"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /clone\_lib="RIKEN full-length enriched, adult male  
 hypothalamus"  
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia



pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with EXassist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

Query Match	26.2%	Score 37.4;	DB 2;	Length 644;
Best Local Similarity	64.4%	Pred. No. 16;		
Matches	56;	Conservative	0;	Mismatches 31; Indels 0; Gaps 0;
29	CCTCTCTCGGCTCGCTCGTTCGGTGGGGGCTCGGACCAAAAGTTCGCAGACGGCAGA	88		
535	CTCTCCGCGCCTGGCCAGGTCGTGGCGCGCGGCGAGACACGACACGCCCATG	594		
89	GCTCTGCTCTGCGGCGCCCAACCGAGCG	115		
595	CTCTACTTTGACGTGCCCGCGGGTG	621		

531460/c	BI531460	604 bp	mRNA	linear	EST 29-AUG-2001
UCUS					
FINITION	1024113H9.y1	C. reinhardtii	CC-1690,	normalized,	Lambda Zap II
	Chlamydomonas	reinhardtii	cDNA,	mRNA sequence.	

BI531460.1 GI:15372034  
EST.  
KEYWORDS

ORGANISM Chlamydomonas reinhardtii  
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;  
Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.  
1 (bases 1 to 604)

**TITLE** Analyses of the Chlamydomonas reinhardtii Genome: A Model, Cellular System for Analyzing Gene Function and Regulation in Vascular Plants. Project: 1024b  
**JOURNAL** Unpublished (2001)

CONTACT: Charles Hauser  
DCMB Box 91000  
Duke University  
Durham, NC 27708-1000  
Tel: 919 613 8159  
Fax: 919 613 8177  
Email: [chauser@duke.edu](mailto:chauser@duke.edu).

```

source
1..604
/organism="Chlamydomonas reinhardtii"
/mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap
II"

```

1.  
/note=Vector: pBluescript II SK-; Site\_1: EcoRI; Site\_2:  
XhoI; This library, constructed by John Davies and Jeffrey  
McDermott, combines cDNAs from CC-1690 cells grown to  
mid-log phase in TAP (acetate-containing) medium in the  
light, TAP medium in the dark, HS (minimal) medium in  
ambient levels of CO<sub>2</sub> and HS medium bubbled with 5% CO<sub>2</sub>.  
PolyA mRNA was purified from each sample, pooled and cDNA  
synthesized. The cDNA was directionally cloned into lambda  
ZAP II (Stratagene) in the EcoRI (5') and XhoI (3') sites.

	Query Match	25.1%	Score 36.2	DB 3	Length 604	
	Best Local Similarity	57.5%	Pred. No. 34			
	Matches 65	Conservative 0	Mismatches 48	Indels 0	Gaps 0	
Qy	16	GAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCCTCGGACCAAGATC	75			
Db	392	GCGTCTCTTCTTCTCTCTGCGCCCAATCTCGTGGATGCGCTTCGGCCCTGCGCGTC	333			
Qy	76	CGCAGACGGCAGAGCTCTGTCTTCGGGGCCCCACCGAGCGAGCGCGCAG	128			
Db	332	GCGGGCGCCAGCGCCCGCTCGCGTGCCTCTTGGCGCTTCGCGAGCTG	280			

CNS0072Q	CNS0072Q	932 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence				
DEFINITION	BACR14B09 of RPCI-98 library from <i>Drosophila melanogaster</i> (fruit fly), genomic survey sequence.				

Accession	AL066742.1	GI:4945205
VERSION	GSS.	
KEYWORDS		
SOURCE	Drosophila melanogaster (fruit fly)	
ORGANISM	Drosophila melanogaster	

REFERENCE  
LITTONS  
Genoscone  
1 (bases 1 to 932)

**JOURNAL**  
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : [secref@genoscope.cns.fr](mailto:secref@genoscope.cns.fr))  
 - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Dr. Jean-Marc Drenth  
 Dr. Jean-Marc Drenth

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila* melanogaster genome using these BACs. For further information please see [http://www.fruitfly.org/TheBDGP/Drosophila\\_melanogaster\\_BAC\\_library](http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library) was prepared by Kazutoyo Osoegawa and Aaron Mammoss in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RP11-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at: [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

```

1. 932
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14B09"
/clone_lib="RPCI-98"
/note="end : T7"
source

```

	Query Match	25.3%	Score 36.2	DB 10	Length 932
	Best Local Similarity	36.3%	Pred. No. 34		
	Matches 41	Conservative 32	Mismatches 40	Indels 0	Gaps 0
Qy	22	CCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGCGCTCGGACCAAAAGTCTCCGAGA	81		
Db	725	SSCGSCCCSCSCCCCCCCCCSSCGSCSSGCGSGCGSCGCGSCGCGSCGCGCGG	784		
Oy	82	CGGCAGAGCTCTGCTCTGCGCGGCCCCCAACCGAGCGAGCGCGCGCAGAGAGGG	134		





```

ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE     1 (bases 1 to 793)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: InCyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM1532 row: e column: 09
              High quality sequence stop: 780.
FEATURES     source
              Location/Qualifiers
                1..793
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5212016"
                /tissue_type="leukocyte"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_118"
                /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
                (destroyed); RNA source leukocytes from anonymous pool of
                non-activated adult donors. Library is oligo-dT primed
                and directionally cloned (EcoRV site is destroyed upon
                cloning). Average insert size 1.7 kb, insert size range
                1.2-3.3 kb. Library is normalized and enriched for
                full-length clones and was constructed by C. Gruber
                (Invitrogen). Research Genetics tracking code 027. Note:
                this is a NIH_MGC Library."
ORIGIN
Query Match      24.9%; Score 35.6; DB 3; Length 793;
Best Local Similarity 55.7%; Pred. No. 49;
Matches 68; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY      7 CTAGTATGAGTTGCCACTCCCTCTCTGCGGGCTGCTCGTGGGGCGCTGCGGA 66
        |||||
Db      641 CGAGTGTGGGACTCTAATCCCATTTGCTCAGCTTAATTTCTGGTAGGTACGCTGA 582

QY      67 CCAAGGTCGCGACGAGAGCTCTGCTCTGCGGGCCCGCCAGCGAGCGGCGC 126
        |||||
Db      581 GCACGGCGCACAGGCCACAGAGCTGTGCACCGCCGGCTCAGGTCCAGGCGAGCTCAC 522

QY      127 AG 128
        ||
Db      521 AG 520

RESULT 11
BM906771/c
LOCUS          BM906771 1115 bp mRNA linear EST 12-MAR-2002
DEFINITION    AGENCOURT 5621680 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5724802
              5', mRNA sequence.
ACCESSION     BM906771
VERSION       BM906771.1 GI:19357150
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominoidea; Homo.
REFERENCE     1 (bases 1 to 1115)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)

ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE     1 (bases 1 to 793)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: InCyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM1532 row: e column: 09
              High quality sequence stop: 780.
FEATURES     source
              Location/Qualifiers
                1..793
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5212016"
                /tissue_type="leukocyte"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_118"
                /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
                (destroyed); RNA source leukocytes from anonymous pool of
                non-activated adult donors. Library is oligo-dT primed
                and directionally cloned (EcoRV site is destroyed upon
                cloning). Average insert size 1.7 kb, insert size range
                1.2-3.3 kb. Library is normalized and enriched for
                full-length clones and was constructed by C. Gruber
                (Invitrogen). Research Genetics tracking code 027. Note:
                this is a NIH_MGC Library."
ORIGIN
Query Match      24.9%; Score 35.6; DB 3; Length 1115;
Best Local Similarity 55.7%; Pred. No. 48;
Matches 68; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY      7 CTAGTATGAGTTGCCACTCCCTCTCTGCGGGCTGCTCGTGGGGCGCTGCGGA 66
        |||||
Db      636 CGAGTGTGGGACTCTAATCCCATTTGCTCAGCTTAATTTCTGGTAGGTACGCTGA 577

QY      67 CCAAGGTCGCGACGAGAGCTCTGCTCTGCGGGCCCGCCAGCGAGCGGCGC 126
        |||||
Db      576 GCACGGCGCACAGGCCACAGAGCTGTGCACCGCCGGCTCAGGTCCAGGCGAGCTCAC 517

QY      127 AG 128
        ||
Db      516 AG 515

RESULT 12
AG080476
LOCUS          AG080476 1123 bp DNA linear GSS 03-NOV-2001
DEFINITION    Pan troglodytes DNA, clone: PTB-076K10.R, genomic survey sequence.
ACCESSION     AG080476
VERSION       AG080476.1 GI:16632278
KEYWORDS      GSS.
SOURCE        Pan troglodytes (chimpanzee)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominoidea; Pan.
REFERENCE     1
AUTHORS      Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
              Tokoki, Y., Watanabe, H. and Sakaki, Y.
TITLE        BAC end sequences of Library PTB
JOURNAL      Unpublished
REFERENCE     2 (bases 1 to 1123)
AUTHORS      Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
              Tokoki, Y., Watanabe, H. and Sakaki, Y.
TITLE        Direct Submission
JOURNAL      Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
              and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
              1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
              (E-mail: chimpansegsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,

```

```

/db_xref="taxon:9598"
/clone="PTB-065P10.F"
/sex="male"
/cell type="lymphoblast"
/clone lib="PTB Chimpanzee Male BAC Library"

```

## ORIGIN

Query Match	24.6%	Score 35.2	DB 10	Length 2024;
Best Local Similarity	59.2%	Pred. No. 61		
Matches 58;	Conservative	0	Mismatches	40
			Indels	0
			Gaps	0

Qy 36 GCGCGCTCGCTCGTTCGGTGGGGCCCTGCCGACCAAGTCCGACACGGCAGAGCTCTGC 95

1047 GCGNCCAGCGCGACGGACGGGACAGGCCGAGAGAGCGGTCGNGAGAGGACGGC 1106

96 TCTGCGGCCCA CCGAGCGAGCGCGCAGAGAGG 133

Db 1107 GCCTCCGGCCGGGCAGAGCGCGCACGCCGGCAGAGAGG 1144

RESULT 14  
CL751558  
724 bp DNA linear CSS 27-III-2004

LOCUS CL/51558 1/24 bp DNA linear GSS 217-001 2000  
DEFINITION OR\_BB0118A16.f OR\_BB0118A16 5',  
accession number

Accession	Version	Genomic survey sequence
CL751558		
CL751558.1	1	CT.50698847

VERSION	01.00000001
KEYWORDS	GSS.
SOURCE	Oruza nijava

SOURCE	ORGANISM
Oryza nivara	Oryza nivara
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;	

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE  
1 (bases 1 to 724)  
Kim, H., Yu, Y., Stun, D., Yost, D., Rao, K., Luo, M., Jetty, R.,  
AUTHORS

TITLE	
Kudrna,D., Muller,C., Hatfield,J., Soderlund,C. and Wing,R.	OMAP Project

**JOURNAL**  
**COMMENT**  
**Unpublished (2004)**  
**Contact: Rod A. Wing**

Arizona Genomics Institute  
University of Arizona

**Forbes Building Room 303, Tucson, AZ 85721-0030, USA**  
**Tel: 520 626 9595**  
**Fax: 520 626 1850**

Fax: 520 621 1259  
Email: [rwing@genome.arizona.edu](mailto:rwing@genome.arizona.edu)

FCK FAKMEIB  
 FORWARD: TAA TAC GAC TCA CTA TAG GG  
 BACKWARD: CAC TCA TTA GCC ACC CCA

Insert Length: 161 Std Error: 0.00  
Plate: 0118 row: A column: 16

Plate: 0110    Lanes: 1-10    Flow: A    Column: 20  
 Seq primer: TAA TAC GAC TCA CTA TAG GG  
 Class: RAC ends.

FEATURES	Location/Qualifiers
source	1..724

```

/organism="Oryza nivara"
/mol type="genomic DNA"

```

```

/db_xref="taxon:4536"
/clone="OR BBa0118A16"

```

```
/tissue type="young leaves"  
/lab host="DH10B-T1 phage resistant"
```

```
/clone_lib="OR_BBa"  
/note="Vector: PAGIBAC1; Site_1: HindIII; Site_2: HindIII"
```

ORIGIN

Query Match	24.5%	Score 35;	DB 10;	Length 724;
Best Local Similarity	55.3%	Pred. No. 70;		

```
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0
```

[illegible][illegible][illegible]



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 18:24:29 ; Search time 5092.96 Seconds  
(without alignments)  
14735.325 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_2829\_4432  
Perfect score: 1604  
Sequence: 1 atgggttcaggcggtggcgc.....tacctaccgcctccctgtga 1604

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_est3:\*
- 4: gb\_est4:\*
- 5: gb\_est5:\*
- 6: gb\_est6:\*
- 7: gb\_est7:\*
- 8: gb\_est8:\*
- 9: gb\_est9:\*
- 10: gb\_est10:\*
- 11: gb\_est11:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	180.8	11.3	264	7	CO892248
C 2	126.8	7.9	195	7	CO888893
C 3	66.8	4.2	753	9	BH115587
C 4	48	3.0	367	3	BJ432209
C 5	47	2.9	581	5	BW335572
C 6	45.6	2.8	681	10	CNS02EOD
C 7	45.6	2.8	889	7	CK159167
C 8	45.4	2.8	581	6	CF569196
C 9	43.6	2.7	414	1	AV964990
C 10	43.4	2.7	801	7	CO465395
C 11	43.2	2.7	748	7	CO048199
C 12	42.8	2.7	415	2	BE556771
C 13	42.8	2.7	668	6	CD920707
C 14	42.2	2.6	435	5	BQ449242
C 15	42.2	2.6	450	3	BJ371297
C 16	42.2	2.6	473	2	BE016852
C 17	42.2	2.6	477	6	CA743784
C 18	42.2	2.6	530	1	AW566595
C 19	42.2	2.6	546	5	BW249566
C 20	42.2	2.6	552	3	BQ285353
C 21	42.2	2.6	660	5	BW193819
C 22	41.8	2.6	539	7	CR929543

23	41.8	2.6	939	10	CNS00CNG
24	41.6	2.6	536	3	BI774214
25	41.6	2.6	599	7	CN788278
26	41.6	2.6	655	7	CK831007
27	41.6	2.6	717	7	CK982139
28	41.4	2.6	620	8	CK180847
29	41.4	2.6	658	5	BX622561
30	41.4	2.6	1116	7	CN068332
31	41.2	2.6	359	3	BI881431
32	41.2	2.6	510	3	BM572584
33	41.2	2.6	546	9	BH589499
34	41.2	2.6	549	3	BM859438
35	41.2	2.6	675	5	BM520599
36	41.2	2.6	713	5	BM534121
C 37	41	2.6	476	1	AL720322
38	41	2.6	612	2	BE195101
39	41	2.6	670	6	CB047069
40	41	2.6	690	7	CK661380
41	41	2.6	720	6	CA296025
42	41	2.6	4479	4	BC044003
43	40.8	2.5	644	3	BM605179
44	40.6	2.5	671	5	BQ987654
45	40.6	2.5	690	5	BQ988583

ALIGNMENTS

RESULT 1  
CO892248/c  
LOCUS  
DEFINITION  
BOVGen.20573 normal cattle brain Bos taurus cdna clone  
RZPDp1056M0360Q 5', mRNA sequence.  
ACCESSION  
CO892248  
VERSION  
CO892248.1  
KEYWORDS  
EST.  
SOURCE  
Bos taurus (cow)  
ORGANISM  
Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
Pecora; Bovidae; Bovinae; Bos.  
REFERENCE  
Hennig S., Janitz M., Herwig R. and Williams J.  
Generation, annotation, evolutionary analysis and database  
integration of 14969 cattle EST clusters  
Unpublished (2004)  
JOURNAL  
Contact: Hennig S  
Laboratory 123, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Inhnstr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1612  
Fax: +49 30 8413 1380  
Email: hennigmolgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting  
(ONFP) to reduce sequencing redundancy. According to the ONFP  
procedure, clones that display the same hybridisation matrix with a  
battery of 200 8mer oligonucleotides are grouped into clusters. One  
clone per ONFP cluster was selected for sequencing. cDNA clones and  
filters are distributed via Deutsches Ressourcenzentrum fuer  
Genomforschung GmbH (<http://www.rzpd.de>).  
PCR Primers  
FORWARD: 5' GGTATACCGCTTACACTTTATGCTTCGGGTCG 3' (M13RSP) 5'-seq  
BACKWARD: 5' GGTATACCGCTTACACTTTATGCTTCGGGTCG 3' (M13FSP) 3'-seq  
Seq primer: 5'-CCGTCGGGATTCCTCCGGGT-3' (M13RSP).  
Location/Qualifiers  
1. 264  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/clone="RZPDp1056M0360Q"  
/sex="female"  
/tissue\_type="brain tissue"  
/dev\_stage="adult brain"

FEATURES  
source



```

/clone lib="normal cattle brain"
/Note="Organ: brain; Vector: pSport1; Site 1: NotI;
Site_2: SalI; Random primed and directionally cloned in
pSport1 vector using NotI
(5'-pGACTAGTTCTAGATCGGAGCGCGGCC (T)15-3' and SalI 5'-
TCGACCCAGCGCTCG-3' adapters (Gibco BRL)"/>
ORIGIN
Query Match 11.3%; Score 180.8; DB 7; Length 264;
Best Local Similarity 80.3%; Pred. No. 5.2e-42;
Matches 212; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1241 GTCCATTGGGCGCAAAATTCCTCACACAGATGGACATTTCCACCGTCTCTCTATGG 1300
DB 264 GTCCATTGGGCGCAAAATTCCTCACACAGATGGACATTTCCACCGTCTCTCTATGG 205
QY 1301 GCGGCTTTGAGCTCAAGAACCGGCTCTCTCACATCTCATCAAAACAGCGCTGTTCCTG 1360
DB 204 GCGGATTGGGACTTAATCAACCTCTCTCACACAGATTTTCATCAAGAACAGCGGTTCTCTG 145
QY 1361 CGAATCTCTCGGGGAGTTTTCAGCTACAAAGTTTGTCTTATTCATCACCCTAATCTCA 1420
DB 144 CGAGTCTCTCGAGCAGCTTCAGTGGCGGAAAGTTTGTCTTATTCATCAGAGTACTCA 85
QY 1421 CAGGACAGTGTGTGGAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGA 1480
DB 84 CCGGTCAAGTCAAGTGGAGATTTGAGTGGAGCTGGGAAAGAAACAGCAAGCGCTGGA 25
QY 1481 ATCCGGAAGTGCAGTACATCA 1504
DB 24 ATCCGGAATTCAGTACACTTCA 1

```

```

RESULT 2
CO888893/c
LOCUS
DEFINITION
BovGen 17218 normal cattle brain Bos taurus cDNA clone
EST 01-SRP-2004
RZPDp105600960Q 5', mRNA sequence.
CO888893
VERSION
KEYWORDS
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
REFERENCE
Hennig S., Janitz M., Herwig R. and Williams J.
Generation, annotation, evolutionary analysis and database
integration of 14969 cattle EST clusters
Unpublished (2004)
JOURNAL
COMMENT
Contact: Hennig S
Laboratory 123, Dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Innestr. 63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1612
Fax: +49 30 8413 1380
Email: hennigmolgen.mpg.de
The library was characterised by oligonucleotide fingerprinting
(ONFP) to reduce sequencing redundancy. According to the ONFP
procedure, clones that display the same hybridisation matrix with a
battery of 200 8mer oligonucleotides are grouped into clusters. One
clone per ONFP cluster was selected for sequencing. cDNA clones and
filters are distributed via Deutsches Ressourcenzentrum fuer
Genomforschung GmbH (http://www.rzpd.de).
PCR Primers
FORWARD: 5' CCCAGCTTTACATTTATGTTCCGGCTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GTATTACCGAGCTGGGAAAGGGGATGTG 3' (M13RSP) 3'-seq
Seq primer: 5'-CCGTCGGAATTCGCCGGT-3' (M13RSP).
FEATURES
Location/Qualifiers
1..195
/organism="Bos taurus"
/mol_type="mRNA"

```

```

/db_xref="taxon:9913"
/clone="RZPDp105600960Q"
/sex="female"
/tissue_type="brain tissue"
/dev_stage="adult brain"
/clone lib="normal cattle brain"
/Note="Organ: brain; Vector: pSport1; Site 1: NotI;
Site_2: SalI; Random primed and directionally cloned in
pSport1 vector using NotI
(5'-pGACTAGTTCTAGATCGGAGCGCGGCC (T)15-3' and SalI 5'-
TCGACCCAGCGCTCG-3' adapters (Gibco BRL)"/>
ORIGIN
Query Match 7.9%; Score 126.8; DB 7; Length 195;
Best Local Similarity 78.4%; Pred. No. 5.8e-26;
Matches 152; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1411 CAATACTCCAGACAGAGTGTGGAATTTGAATGGAGCTGCAGAAAGAAACAGC 1470
DB 194 CAGTACTCCAGGACAGGTGAGGTGGAGATCAGTGGGAGTTGCAGAGGTAACAGC 135
QY 1471 AAGCGTGGAAATCCGAAGTGCAGTACATCCAAATTTATGCAAAATTCGCAAGTTGAT 1530
DB 134 AAACGGTGGAAATCCGAATTTCACTACACTTCCAACTACAACAAGTCTGTAAATGTGAT 75
QY 1531 TTTACTGTGGACAACTATGGACTTTTATCTAGCTCGCCCATTTGGCACCGTTACCTT 1590
DB 74 TTTAATGTGGACATTTAATGGCGTGTATTTCAGAGCTCGCCCATTTGGCACGATACCTG 15
QY 1591 ACCGTCCTCCCTGTA 1604
DB 14 ACTCGTAATGTGTA 1

RESULT 3
BH115587
LOCUS
DEFINITION
RPCI-24-358F16.TV RPCI-24 Mus musculus genomic clone
RPCI-24-358F16, genomic survey sequence.
BH115587
ACCESSION
VERSION
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidae; Muridae; Murinae; Mus.
REFERENCE
Zhao S., Nierman W., Malek J., Shatman S., Akinret B., Levins M.,
Tsagaye G., Geer K., Krol M., Shvartsbeyn A., Gebregorgis E.,
Russell D., de Jong P. and Fraser C.M.
1 (bases 1 to 753)
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other GSRS: RPCI-24-358F16.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
page: http://www.tigr.org/db/bac_ends/mouse/bac_end_intro.html
Plate: 358 row: F column: 16
Seq primer: 17
Class: BAC ends.
Location/Qualifiers
1..753
/organism="Mus musculus"
/mol_type="genomic DNA"

```

FEATURES  
source

	Matches	120;	Conservative	0;	Mismatches	122;	Indels	0;	Gaps	0;
Qy	110	TCATCA	CA	CGACG	ACCTTGGGGCCTTGGCCACCTACATAAACCACTCTTACAAGC	169				
Db	328	TCAACAACAT	CANCAACAT	CAANTACCTCTTAAACCAACCTACATCAACAACATCAACT	269					

```

Db      388 ACAACAGGACCAACACACACAGGAAACAACAGCACACAGCAATCATCAATA 447
      383 TCACACACCTCGTATACCTTACACAGCAGGTTCAAGTCTTCGCGACT 433
      448 ACAACACACGCTTAATACAGAAACAACACAAACACATCACCACCACT 498

RESULT 6
CNS02EOD/LOCUS      681 bp DNA linear GSS 01-SEP-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
              262H14 of library G from Tetraodon nigroviridis, genomic survey
              sequence.
ACCESSION AL193990
VERSION AL193990.1 GI:7932096
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
          Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
          Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1
          Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
          Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
          Saurin,W. and Weissenbach,J.
          Estimate of human gene number provided by genome-wide analysis
          using Tetraodon nigroviridis DNA sequence
          Nat. Genet. 25 (2), 235-238 (2000)
          10835645
REFERENCE 2
          Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
          Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
          Saurin,W., Bernot,A. and Weissenbach,J.
          Characterization and repeat analysis of the compact genome of the
          freshwater pufferfish Tetraodon nigroviridis
          Genome Res. 10 (7), 939-949 (2000)
          10899143
REFERENCE 3 (bases 1 to 681)
          Genoscope.
          Direct Submission
          Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
          BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr]
          Web : www.genoscope.cns.fr]
COMMENT This sequence is a single read and was generated as part of a large
          scale clone-end sequencing project of the Tetraodon nigroviridis
          genome. For more information, please take a look at
          http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
          source
            1..681
              /organism="Tetraodon nigroviridis"
              /mol_type="genomic DNA"
              /db_xref="taxon:99883"
              /clone="262H14"
              /clone_lib="G"
              /note="Genoscope sequence ID : C0AG262DD07LPI
              end : T7"
ORIGIN
Query Match      2.8%; Score 45.6; DB 10; Length 681;
Best Local Similarity 43.4%; Pred. No. 0.15;
Matches 144; Conservative 10; Mismatches 178; Indels 0; Gaps 0;

QY      31 GACAATAACGAAGCGCCGACGAGTGGTAAATGCTCAGGAAATTTGGCATTTCC 90
      351 GACAACACGACACACACACACACACACACACACACACACACACACACACAC 292

QY      91 ACATGGCTGGGCGACAGAGTATCATCCACGACCGCACTGGCCCTTGGCCACTTAC 150
      291 RACGGCAGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 232

QY      151 AATAACCACTCTTACAAGCAAAATCTCCAGTGTCTTCAACGGGGCCGACGACCAACCAC 210

```

```

Db      231 AACACACACACACACACACACACACACACACACACACACACACACACAC 172
      211 TACTTTCGGCTACAGCACCCCTCGGGGTATTTTGATTTTCAACAGATTTCACCTTTT 270
      171 AAAAAACACACACACACACACACACACACACACACACACACACACACACACAC 112
      271 TCACACGTGACTGCAGGACTCATCAACACAAATTGGGGATTTCGGGCCAAGAGATC 330
      111 AACAAAAACACACACACACACACACACACACACACACACACACACACACAC 52
      331 AACTTCAACTCTTCAACATCAAGTCAAGGA 362
      51 AASMACACACACACACACACACACACACACACACACACACACACACACAC 20

CK159167      869 bp mRNA linear EST 05-DEC-2003
FGAS040564 Triticum aestivum FGAS: TaIt5 Triticum aestivum cDNA,
mRNA sequence.
CK159167
CK159167.1 GI:38985053
EST.
Triticum aestivum (bread wheat)
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 869)
          Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,
          Geneswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,
          Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilsson,D.,
          Pennikett,C., Roach,J.L. and Sarhan,F.
          Functional Genomics of Abiotic Stress In Wheat and Canola Crops
          Unpublished (2003)
          Contact: Wm L Crosby
          Bioinformatics
          University of Saskatchewan, Department of Computer Science
          1C101 Engineering Building, 57 Campus Drive, Saskatoon,
          Saskatchewan, S7N 5A9, Canada
          Tel: 306 966 1769
          Fax: 306 966 2033
          Email: fgas_ests@cs.usask.ca
          This sequence is the direct result of the Base calling software
          Phred (default parameters). It is the raw base calls. To aid in the
          identification of the high quality insert the software Lucy
          (default parameters) has been run on this sequence. Lucy identified
          the region [128,636].
          Plate: TaIt537 row: N column: 23.
FEATURES             Location/Qualifiers
          source
            1..869
              /organism="Triticum aestivum"
              /mol_type="mRNA"
              /cultivar="Wheat line PI 178383"
              /db_xref="taxon:4565"
              /lab_host="DH5 alpha"
              /clone_lib="Triticum aestivum FGAS: TaIt5"
              /note="Organ: Crown; Vector: pGEM-T; SSH (suppression
              subtractive hybridization) cDNA library from genotype
              PI178383 cold hardened at 2 C for 21 days and 49 days
              (equal amount of cDNA pooled together before subtraction,
              tester) and subtracted against genotype Norstar cold
              hardened at 2 C for 1 day (24 H)(driver). Modified Smart
              cDNA (Clontech)priming and non-directional cloning"
ORIGIN
Query Match      2.8%; Score 45.6; DB 7; Length 869;
Best Local Similarity 48.1%; Pred. No. 0.17;
Matches 129; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY      145 ACCTACAATAACCACTCTTACAAGCAAAATCTCCAGTGTCTTCAACGGGGCCGACGACGAC 204

```

Db	785	AACAAACCAACCAACCAACAAACGACGACGACGACAAACGACCAACCAACCAAC	726
Qy	205	AACCACTACTTCGGCTACAGCACCCCTCGGGGTATTTTGATTTCAACAGATTCACATGC	264
Db	725	AACAACCAACAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	666
Qy	265	CACITTTTACCACGTCGAGCTGCGCAGGACTCATCAACAACATTTGGGGATTTCGGGCCAAG	324
Db	665	CACCAACCAACAAACAAACAAACAAACGACAAACAAACAAACAAACAAACAAAC	606
Qy	325	AGACTCAACTTCAAACTCTTCAACATCCAAGTCAAGGAGGTCAACGAAATGATGGCGTC	384
Db	605	AACAACCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC	546
Qy	385	ACAACCATCGCTAATAACCTTACCACGA	412
Db	545	AACAACCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC	518

RESULT 8	CF569196	581 bp	mrna	linear	EST 08-SEP-2004
LOCUS	CF569196				
DEFINITION	EST057 Subtracted, Clontech (cat. # K1804-1) Triticum aestivum cDNA clone FDC57 5', mRNA sequence.				
ACCESSION	CF569196				
VERSION	CF569196.1	GI:51921536			
KEYWORDS	EST.				
SOURCE	Triticum aestivum (bread wheat)				
ORGANISM	Triticum aestivum				
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.				
REFERENCE	1. (bases 1 to 581)				
AUTHORS	Xiao,K., Bai,G.H. and Carver,B.F.				
TITLE	Nylon Filter Arrays Reveal Differential Expression of Expressed Sequence Tags in Wheat Roots Under Aluminum Stress				
JOURNAL	J. Integr. Plant Biol. 47 (7), 839-848 (2005)				

```

FEATURES
source
High quality sequence stop: 581.
Location/Qualifiers
1..581
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="OK91G106"
/db_xref="taxon:4565"
/clone="FDC57"
/tissue_type="root"
/clone_lib="Subtracted, Clontech (cat. # K1804-1)"
/note="EST from wheat ('Triticum aestivum, cv. OK91G106)
root in response to aluminum stress"

```

ORIGIN	Query Match	2.8%	Score 45.4	DB 6	Length 581
	Best Local Similarity	46.6%	Pred. No. 0.16		
	Matches 145	Conservative 0	Mismatches 166	Indels 0	Gaps 0
Qy	102	CGACAGAGTCATACCCACACGACCCCGCACTCTGGGCTTGGCCACCTACCAATRACCACCT	161		
Db	152	CAACAACAACAACAACCAACCAACACAGCAGCAACAGCAACAACAGCAACAACAA	211		
Qy	162	CTACAAGCAAAATCTCCAGTGTCTCAAACGGGGGCCAGCAACGCACCACTACTTTCGGGCTA	221		
Db	212	CAACAACAACAACAGCAACAACACAGCAACAACCAACAACAACATCAGCAACAACAA	271		
Qy	222	CAGCACCCCTCGGGGGTATTTTGATTTTCAAACAGATTTCCATTCGCGCACTTTTCCACCGTGA	281		
Db	272	CAGCAGACCAACACAGCAGCAACAACACAGCAGCAGCAGCAACAGCAACAGTAACAACAGCAAC	331		

Qy	282	CTGGCAGC	ACTCATCA	CAACAATT	GGGATTC	CGGCCCA	GAGAGACT	CAACTTCA	AACT	341
Db	332	CAACAGC	AGCAAC	CAACCA	CAACCA	CAACCA	CAACCA	CAACCA	CAACAA	391
Qy	342	CTTCAAC	ATCCAG	TCAAGAG	GTCAGG	AGCATG	TGGCGT	CACACC	ATCGCT	401
Db	392	CAGCAAC	CAACAA	CAACCA	CAACAT	CAGCAAC	CAACAT	CAGCAAC	CAACAG	451
Qy	402	CCTTACC	GACGA	412						
Db	452	CAGCAAC	GACGA	462						

RESULT 9	AV964990/c	414 bp	mRNA	linear	EST 14-MAR-2002
LOCUS	AV964990	Nozi	Satoh unpublished	cdna library,	young adult Ciona
DEFINITION		intestinalis	cdna clone ciad17116	5',	mRNA sequence

Accession	AF047826
Version	AV964990.1
Keywords	GI:19454686
Source	EST.
Organism	Ciona intestinalis
Reference	Ciona intestinalis
Authors	Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Clonidae; Ciona.
Title	1 (bases 1 to 414)
Journal	Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.
Comment	Expressed genes in Ciona intestinalis Unpublished (2000) Contact: Nori Satoh

Kyoto University  
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan  
Tel: 81-75-753-4081  
Fax: 81-75-705-1113  
Email: satohe@ascidian.zool.kyoto-u.ac.jp.

## FEATURES

```
source
organism="Clona intestinalis"
no_type="rRNA"
db_xref="taxon.7719"
clone="ciad1716"
tissue_type="whole animal"
dev_stage="young adult"
clone_lib="Nori satoh unpubli
adult"
```

## ORIGIN

Query Match	2.7%	Score 43.6	DB 1	Length 414
Best Local Similarity	46.6%	Pred. No. 0.49		
Matches 139	Conservative	0	Mismatches 159	Indels 0
Gaps 0				

Qy	111	CATCACACAGACACCGCAGCTTGGGGCTTGGCCACTACAAATTAACAACCTTCTACAGCA	170
Db	320	CAACAAACAACAACCTGTTCTACAGGCACAACAACAACCACTGATTCAACAACAACAACAA	261
Qy	171	AATCTCAGACTGCTTCAACGGGGGCAGCAACGACAACCACTACTTCGGGTACAGCACCCCC	230
Db	260	CNACAACTGTTGCTACAGGCACAACGACACCAACCAACCTTAATTCACAACAACAACAACAA	201
Qy	231	CTGGGGGTATTTGATTTTAAACAGATTCGACTGCGCACTTTTCCACCGTGACTGGGAGCG	290
Db	200	CTGTTGCTACAGGCACAACAACAACAACAACAACCTCTGATTCAACAACAACAACAACAACCTG	141
Qy	291	ACTCATCAACAACAATTTGGGGATTCCGGGCCAAGAGAGACTCAACTTTCAAACTCTTCAACAT	350
Db	140	TTGCAGCGGCAACAACAACGACACAACAACAACCTGATTTCACAACAACAATTTGCTACAG	81
Qy	351	CCAAGTCAAGGAGGTCAAGCAATGATGGGTCAACAACCATCGCTAATTAACCTTACC	408
Db	80	GCACAAACAACAACAACAACAACAACCTGTTGCAACAGGAAACAACAGCAACAACAACAACCC	23

RESULT 10	CO465395	801 bp	mRNA	linear	EST 08-JUN-2005			
LOCUS	MZCCS20039C06.g	Maize Endosperm	cdna	Library	Zea mays cdna, mRNA			
DEFINITION	sequence.							
ACCESSION	CO465395							
VERSION	CO465395.1	GI:67039140						
KEYWORDS	EST.							
SOURCE	Zea mays							
ORGANISM	Zea mays							
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD Clade; Panicoideae; Andropogoneae; Zea.							
REFERENCE	1 (bases 1 to 801)							
AUTHORS	Verza,N.C., Silva,T.R., Cord-Neto,G., Nogueira,F.T.S., Fisch,P.H., De Rosa,V.E. Jr., Martins,M.M., Vettore,A.L., da Silva,P.R. and Arruda,P.							
TITLE	Endosperm-preferred expression of maize genes as revealed by transcriptome-wide analysis of expressed sequence tags							
JOURNAL	Plant Mol. Biol. (2005) in press							
COMMENT	Contact: Arruda P Centro de Biologia Molecular e Engenharia Genetica Universidade Estadual de Campinas Caixa Postal 6010, 13083-970, Campinas SP, Brazil Tel: 55 19 3788 1137 Fax: 55 19 3788 1089 Email: parruda@unicamp.br.							
FEATURES	Location/Qualifiers							
source	1..801							
	/organism="Zea mays"							
	/mol_type="mRNA"							
	/cultivar="F-352 near isogenic line"							
	/db_xref="taxon:4577"							
	/sex="hermaphrodite"							
	/tissue_type="endosperm"							
	/dev_stage="multiple stages (10 to 25 days after pollination. see publication for more information)"							
	/lab_host="E. coli DH10S"							
	/clone_lib="Maize Endosperm cdna Library"							
	/notes="Organ: seed; Vector: pSPORT1; Site: 1: SalI; Site:2: NotI; Plant Material and RNA Isolation: Field grown maize plants from inbred line F352 were used. Ears were harvested at 10, 15, 20 and 25 days after pollination (DAP), seeds were cut from the ear and the upper third of the endosperms, containing only endosperm, aleurone and pericarpal tissues, was removed, frozen in liquid nitrogen and stored at -80o C. Frozen endosperms were pulverized in liquid nitrogen and total RNA was isolated according the method of Manning (9). Poly(A)+RNA was isolated using Oligotex-dT. cdna libraries were constructed using Superscript Plasmid System for cDNA Synthesis and Plasmid Cloning Kit as described in Vettore, et al., (2001). The libraries that made SUCEST. Genet Mol Biol 24: 1-7. cDNAs ranging from 500 to 800 bp in size were assigned as short libraries (S10, S15, S20), and cDNAs >800 were assigned as long libraries (L10, L15, L20, L25). Unamplified libraries were plated and individual colonies picked and transferred to 96 well plates containing liquid Circie Grow (CG) medium supplemented with 100 mg/L of ampicillin and 8% glycerol. Three copies of each cdna clone were stored at -80o C. Additional information can be found in : Verza, N.C., Silva, T.R., Cord-Neto, G., Nogueira, F.T.S., Fisch, P.H., De Rosa Jr., V.E., Martins, M.M., Vettore, A.L., da Silva, P.R. and Arruda, P. (2005) Endosperm-preferred expression of maize genes as revealed by transcriptome-wide analysis of expressed sequence tags. Plant Molecular Biology (/in press/)"							

ORIGIN	FRANK MOLECULAR BIOLOGY (in press)			
Query Match	2.7%	Score 43.4	DB 7	Length 801
Best Local Similarity	46.3%	Pred. No. 0.73		
Matches 143	Conservative 0	Mismatches 166	Indels 0	Gaps 0

QY	CGACAGAGTATCATCACACACGACCCGCACTTGGGGCTTGGCCCACTTACAAATTAACCACTT	161
Db		
162	CTACAGCAAAATCTCAGTGTCTTCAACGGGGGGCCAGCAAGCAACCACTACTTTCGGCTA	221
Db		
122	CAACCAACCAACCAACATCATCCACAACAACATCATCCACAACCAACCAACATCA	181
Db		
222	CAGCACCCCTCGGGGTATTTTGATTTTCAACAGATTCCACTGCGCACTTTTCCACCACTGA	281
Db		
182	CAATCAACAGGCCATGTGGAAACACAGAAACAGAGCATGAGCAACACCAAGGACAGAGC	241
Db		
282	CTGGCAGCAGACTCATCAACAAACAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACT	341
Db		
242	CATGAGCAACCAACATCAACAACAATTTCCAGGGTCAATGACAAGCAGCAACAACCAACAG	301
Db		
342	CTTCAACATCCAAGTCAAGAGGTCAGAGGAATGATGGGTCAACAACCATCGCTAATAA	401
Db		
302	CCTCAGCAATATCAGCAGGGCCAGGAAAAATCAACAGCAACAATGTCTATTGCCAGGAG	361
QY	402 CCTTACCAG 410	
Db		
362	CAGCAACAG 370	
LOCUS	CO048199	
DEFINITION	CO048199.1 GI:48588353	
ACCESSION	EST.	
VERSION	Lumbricus rubellus (humus earthworm)	
KEYWORDS	Lumbricus rubellus	
SOURCE	Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;	
ORGANISM	Lumbricina; Lumbricidae; Lumbricus.	
REFERENCE	1 (bases 1 to 748)	
AUTHORS	Chaseley, J., Hedley, B.A., Morgan, J.C., Sturzenbaum, S., Kille, P. and Blaxter, M.	
TITLE	The Lumbricus rubellus EST program - Sequences from a Fluorantene Exposure Library	
JOURNAL	Unpublished (2003)	
COMMENT	Contact: Jennifer Chaseley	
	BIOS 1	
	Cardiff University	
	Main College, Museum Avenue, Cardiff, CF11 3PL, UK	
	Tel: +44 2920876680	
	Fax: +44 2920874305	
	Email: chaseley@cardiff.ac.uk, Kille@cardiff.ac.uk	
	PCR Primers	
	FORWARD: gtttcccgctcagcagc	
	BACKWARD: caggaacagctatgaccatg	
	Plate: 66 row: A column: 08	
	Seq primer: SKplus	
	High quality sequence start: 8	
	High quality sequence stop: 524.	
FEATURES	Location/Qualifiers	
source	1..748	
	/organism="Lumbricus rubellus"	
	/mol_type="mRNA"	
	/db_xref="taxon:35632"	
	/clone="Lr PAHCF 66A08"	
	/tissue_type="Whole worm"	
	/dev_stage="Adult"	
	/clone_lib="Earthworm Fluorantene Exposure Library"	
	/notes="Vector: pBluescript II SK+; The library was prepared using protocols given by the supplier (Stratagene)."	
ORIGIN		
Query Match	2.7%; Score 43.2; DB 7; Length 748;	
Best Local Similarity	51.6%; Pred. No. 0.62;	

Matches	99;	Conservative	0;	Mismatches	93;	Indels	0;	Gaps	0;
114	y	CACCA	CAGCAGCCCG	CACCTGGG	CGCTTGGCC	CACCTCAAT	AACCACTTCA	CAGCAAA	173
79	b	CGCCC	CACAACCA	CACTACA	CTACTCT	CACCTCTT	CACACGCC	CACACAACCA	138
174	y	CTC	CAGTCTTCA	CGGGGG	CCGCAACG	CAACCACTTTC	GGCTACAG	CACCCCTG	233
139	b	CTACA	CACTACA	CTCTCT	TACCA	CACCAACCA	CACTACA	CACTCACTC	198
234	y	GGGGT	ATTTGATTTCA	ACAGATTCCA	CTGGC	CACTTTTCA	CCACGTCAGT	CGGCGGACT	293
199	b	CGACT	TACCACG	CCCAACCA	CCACTACA	ACTACTCT	CACTCCGACT	TACCACGCCA	258
294	y	CATCA	CAACAA	305					
259		CAACA	CAACCA	270					

EST

RESULT 12  
BE556771  
OCUS  
DEFINITION  
Accession BE556771 linear EST 30-AUG-2000  
fk95c07.y1 Zebrafish Research Genetics C32 fin Danio rerio cdna 5'  
similar to contains element TAR1 repetitive element ; mRNA  
sequence.  
Accession BE556771 415 bp  
VERSION BE556771.1 GI:9821261  
KEYWORDS EST.  
SOURCE Danio rerio (zebrafish)  
ORGANISM Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 415)  
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M.,  
Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,  
Underwood,K., Steptoe,M., Thesling,B., Allen,M., Bowers,Y.,  
Perron,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,  
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,  
Waterston,R. and Wilson.R.  
WashU Zebrafish EST Project 1998  
Unpublished (1998)  
Contact: Stephen L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: zbrfish@watson.wustl.edu  
cdNA Library Preparation: Ning Wu. cdNA Library Arrayed by:  
Research Genetics. DNA Sequencing by: Washington University Genome  
Resequencing Center. Clone distribution: Research Genetics web  
address: <http://www.researchgenetics.com/>  
Seq primer: T3 ET from Amersham  
High quality sequence stop: 401.

```

FEATURES
source
Location/Qualifiers
1. .415
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/tissue_type="fin"
/lab_host="Genethogs (HS996, a phage-resistant isolate of
DH10B)"
/clone_lib="zebrafish Research Genetics C32 fin"
/note="vector: pT73D-Pac with a modified polylinker;
Site 1: EcoRI; Site 2: NotI; 1st strand cDNA was prepared
from zebrafish(C32) fin, and was then primed with a Not I
- oligo(4T) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library is non-normalized. Library was
constructed by Ning Wu. NOTE: This clone is available
royalty-free through LLNL; contact the IMAGE Consortium
(info.llnl.gov) for further information"

```

ORIGIN		Query Match	2.7%	Score 42.8;	DB 2;	Length 415;
		Best Local Similarity	54.4%;	Pred. No. 0.86;		
		Matches	86; Conservative	0; Mismatches	72; Indels	0; Gaps
Qy	111	CATCACACGACACGGCAGCTTGGGCTTGGCCACCTACAATAACCACTCTACAA	GCA	170		
Db	41	CATCAGCACTACACCTACAACCTGCATCTTCATCACTACACCTACAACCTGCGGCTTCAG		100		
Qy	171	AATCTCOAGTGTTCACCGGGGGCGACGAACGCAACCACTACTTGGCTTACGACACCCC		230		
Db	101	CAACTACACCTACAACTGCGGGCTTCAGCACTACACCCACAACCTGCGGCTTCAGCACTA		160		
Qy	231	CTGGGGGTATTTGATTTCACAGATTCCACTGCCACT		268		
Db	161	CACCTTCAACTGGGGCTTCAGCAACTACACCTTCAACT		198		

RESULT 13	
CD920707	
LOCUS	668 bp mRNA linear EST 15-JUL-2003
DEFINITION	G508.118C06F010910 G608 Triticum aestivum cDNA clone G608118C06, mRNA sequence.
ACCESSION	CD920707
VERSION	CD920707.1 GI:32768471
KEYWORDS	EST.
SOURCE	Triticum aestivum (bread wheat)
ORGANISM	Triticum aestivum
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.
AUTHORS	1 (bases 1 to 668)
TITLE	Genoplate.
COMMENT	Genoplate, a major partnership french program in plant genomics Unpublished (2003)
	Contact: Genoplate
	Genoplate
	93, rue Henri Rochefort 91025 EVRY CEDEX France
	Tel: 33 1 69 47 54 00
	Fax: 33 1 69 47 54 10
	This sequence has been generated in the framework of the french plant genomics programme 'Genoplate' ( <a href="http://www.genoplate.com">http://www.genoplate.com</a> and <a href="http://genoplate-info.infobiogen.fr">http://genoplate-info.infobiogen.fr</a> ).

```

FEATURES
source
and http://genome.princeton.edu/pombe/annotation/277.
Location/Qualifiers
1. 668
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="G60818C06"
/tissue_type="grain (608 degrees per day after
pollination)"
/clone_lib="G608"

```

ORIGIN					
	Query Match	2.7%; Score 42.8; DB 6;	Length 668;		
	Best Local Similarity	49.5%; Pred. No. 1;			
	Matches 110; Conservative	0; Mismatches 112; Indels	0; Gaps	0	
Qy	84	CGATTCCACATGGCTGGGCGCAGAGTCATCACCAACGACGCCGCACCTGGGGCCTTGCC	143		
Dd	302	CAAAATCCCCHGCAACCACACAATTTCCCCCAACCCNACACCAACCCACACACCCAC	361		
Qy	144	CACCTTACAATAACCACTCTTAACAGCAAATTCCTCAGTGCTTTCACGGGGGGCAGCAACGA	203		
Dd	362	CACCACCAATTAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA	421		
Qy	204	CAACCACTACTTCGGCTACAGCAACCCCTGGGGGTATTTTCATTTTCAACAGATTTCCA	263		
Dd	422	CAACACAGCAACCAACCAACAGCAACAACAGCAACAACAGCAACAGCAACAGCAA	481		
Qy	264	CCACTTTTTCCACGTCGACTGGCAGCGGACTCATCAACAACAA	305		







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 19:15:02 ; Search time 228.137 Seconds  
(without alignments)  
12497.813 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_2829\_4432

Perfect score: 1604  
Sequence: 1 atggcttcaggcggtggcgc.....taccttaccgtccctgtga 1604

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PTUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/BACKfileseq1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1604	100.0	1605	US-09-807-802A-16	Sequence 16, Appl
2	1604	100.0	1800	US-09-807-802A-14	Sequence 14, Appl
3	1604	100.0	2211	US-09-807-802A-12	Sequence 12, Appl
4	1604	100.0	4718	US-09-807-802A-1	Sequence 1, Appl
5	1604	100.0	7447	US-10-216-870-11	Sequence 11, Appl
6	1512.8	94.3	4683	US-09-807-802A-19	Sequence 19, Appl
7	1026.4	64.0	4072	US-09-770-315-4	Sequence 4, Appl
8	1026.4	64.0	4679	US-10-038-972A-12	Sequence 12, Appl
9	1026.4	64.0	7557	US-09-770-315-3	Sequence 3, Appl
10	1026.4	64.0	8698	US-09-770-315-2	Sequence 2, Appl
11	1021.4	63.7	8179	US-09-438-268-5	Sequence 5, Appl
12	1004.2	62.6	4680	US-08-254-358-1	Sequence 1, Appl
13	1004.2	62.6	4680	US-08-475-391-1	Sequence 1, Appl
14	1004.2	62.6	4680	US-08-709-609-1	Sequence 1, Appl
15	1004.2	62.6	4680	PCT-US95-07178-1	Sequence 1, Appl
16	999.6	62.3	4681	US-09-807-802A-18	Sequence 18, Appl
17	987	61.5	4675	US-09-782-378A-1	Sequence 1, Appl
18	987	61.5	4675	US-09-782-378A-2	Sequence 2, Appl
19	902.8	56.3	8151	US-09-438-268-2	Sequence 1, Appl
20	588.2	36.7	7214	US-09-438-268-1	Sequence 1, Appl
21	588.2	36.7	7744	US-10-216-870-14	Sequence 14, Appl
22	587.2	36.6	1617	US-09-532-594B-19	Sequence 19, Appl
23	587.2	36.6	1800	US-09-532-594B-17	Sequence 17, Appl
24	587.2	36.6	2208	US-09-532-594B-5	Sequence 5, Appl

25	587.2	36.6	4767	3	US-09-532-594B-1	Sequence 1, Appl
26	537.4	33.5	2264	3	US-09-533-427-8	Sequence 8, Appl
27	537.4	33.5	2264	3	US-09-533-427-9	Sequence 9, Appl
28	537.4	33.5	2307	3	US-09-533-427-7	Sequence 7, Appl
29	537.4	33.5	4652	3	US-09-533-427-1	Sequence 1, Appl
30	46	2.9	5511	3	US-08-928-361B-2	Sequence 2, Appl
31	46	2.9	5511	3	US-09-588-995A-2	Sequence 2, Appl
32	46	2.9	7334	3	US-08-928-361B-1	Sequence 1, Appl
33	46	2.9	7334	3	US-09-588-995A-1	Sequence 1, Appl
34	42.8	2.7	5163	3	US-08-700-651-1	Sequence 4, Appl
35	42.8	2.7	5163	3	US-08-928-361B-4	Sequence 4, Appl
36	42.8	2.7	5163	3	US-09-588-995A-4	Sequence 4, Appl
37	42.8	2.7	5318	3	US-08-700-651-2	Sequence 2, Appl
38	42.8	2.7	5318	3	US-08-928-361B-3	Sequence 3, Appl
39	42.8	2.7	5318	3	US-09-588-995A-3	Sequence 3, Appl
40	41.4	2.6	1723	3	US-10-104-047-814	Sequence 814, App
C 41	41	2.6	34662	3	US-09-902-540-1261	Sequence 1261, Ap
C 42	40.4	2.5	601	3	US-09-949-016-165983	Sequence 165983,
43	40.4	2.5	61461	3	US-09-949-016-16419	Sequence 16419, A
44	39.4	2.5	2271	3	US-09-438-268-3	Sequence 3, Appl
45	39.2	2.4	700	3	US-10-187-253E-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-807-802A-16  
; Sequence 16, Application US/09807802A  
; Patent No. 6759237  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,  
; TITLE OF INVENTION: Vectors and Host Cells Containing Same  
; FILE REFERENCE: GNVN.031USA  
; CURRENT APPLICATION NUMBER: US/09/807,802A  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/107,114  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: PCT/US99/25694  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 1605  
; TYPE: DNA  
; ORGANISM: AAV-1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1602)  
; OTHER INFORMATION:  
US-09-807-802A-16

Query Match 100.0%; Score 1604; DB 3; Length 1605;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCGCAGGTGGGT	60
DB	1	ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCGCAGGTGGGT	60
QY	61	AATGCTTCAGGAAATTTGGCATTTCCATGCTGGGCGACAGAGTCATCACACC	120
DB	61	AATGCTTCAGGAAATTTGGCATTTCCATGCTGGGCGACAGAGTCATCACACC	120
QY	121	AGCACCCGACCTGGGCGCTTGGCCACCTCAATAACACCTCTTACAGCAAAATCTCCAGT	180
DB	121	AGCACCCGACCTGGGCGCTTGGCCACCTCAATAACACCTCTTACAGCAAAATCTCCAGT	180
QY	181	GCTTCAACGGGGCCAGCAACGACCACTACTTGGCTACAGACCCCTCGGGGTAT	240
DB	181	GCTTCAACGGGGCCAGCAACGACCACTACTTGGCTACAGACCCCTCGGGGTAT	240

```
QY 241 TTTGATTTCAACAGATTCACACTGACCTTTTACACAGCTGACTGGCAGGACTCATCAAC 300
Db 241 TTTGATTTCAACAGATTCACACTGACCTTTTACACAGCTGACTGGCAGGACTCATCAAC 300
QY 301 AACAAATTGGGGATTCGGGCCCAAGAGACTCAACTTTCAAACCTCTTCAAACATCCAAGTCAAG 360
Db 301 AACAAATTGGGGATTCGGGCCCAAGAGACTCAACTTTCAAACCTCTTCAAACATCCAAGTCAAG 360
QY 361 GAGGTCAACGAAATGATGGGTCAACACCATCGGTAAATAACCTTTACCAAGCAAGGTTCAA 420
Db 361 GAGGTCAACGAAATGATGGGTCAACACCATCGGTAAATAACCTTTACCAAGCAAGGTTCAA 420
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTAGCTCTCGGTCTGCGGACACGAGGCTGC 480
Db 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTAGCTCTCGGTCTGCGGACACGAGGCTGC 480
QY 481 CTCCCTCCGTTCCCGCGGACCGTGTTCATGATTCGCGCAATACGGCTACCTGACGCTCAAC 540
Db 481 CTCCCTCCGTTCCCGCGGACCGTGTTCATGATTCGCGCAATACGGCTACCTGACGCTCAAC 540
QY 541 AATGGCAGCAAGCCGTGGGACGTTTCATCTTTTACTGCTCGAAATATTTTCCCTTCTCAG 600
Db 541 AATGGCAGCAAGCCGTGGGACGTTTCATCTTTTACTGCTCGAAATATTTTCCCTTCTCAG 600
QY 601 ATGCTGAGAACCGGCACAACTTTACCTTCAGTACACCTTTGAGGAAGTGCCTTCCAC 660
Db 601 ATGCTGAGAACCGGCACAACTTTACCTTCAGTACACCTTTGAGGAAGTGCCTTCCAC 660
QY 661 AGCAGCTACCGCACAGCAGCAGCGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720
Db 661 AGCAGCTACCGCACAGCAGCAGCGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720
QY 721 CTGTATTACTCGAACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAACAAAGACTTGTGTG 780
Db 721 CTGTATTACTCGAACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAACAAAGACTTGTGTG 780
QY 781 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACTCGGACCC 840
Db 781 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACTCGGACCC 840
QY 841 TGTATTCCGCGAGCGCGCTTTCTAAAAACAAAAACAGACAAACAAACAGCAATTTTACC 900
Db 841 TGTATTCCGCGAGCGCGCTTTCTAAAAACAAAAACAGACAAACAAACAGCAATTTTACC 900
QY 901 TGGACTGGTCTTCAAAATATATAACCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 960
Db 901 TGGACTGGTCTTCAAAATATATAACCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 960
QY 961 GCTATGGCTCACAAAGACGACGAGCAAGTCTTTTCCCATGAGCGGTGTCATGATT 1020
Db 961 GCTATGGCTCACAAAGACGACGAGCAAGTCTTTTCCCATGAGCGGTGTCATGATT 1020
QY 1021 TTTGGAAAAGAGAGCGCCGGAGCTTCAAAACATGTCATTTGGACAATGTCTATGATTACAGAC 1080
Db 1021 TTTGGAAAAGAGAGCGCCGGAGCTTCAAAACATGTCATTTGGACAATGTCTATGATTACAGAC 1080
QY 1081 GAAGAGGAAATTAAGGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140
Db 1081 GAAGAGGAAATTAAGGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140
QY 1141 AATTTCCAGACGACGACGACGACCTTGGACCGGAGATGTGATCTATGGGAGCAATTA 1200
Db 1141 AATTTCCAGACGACGACGACGACCTTGGACCGGAGATGTGATCTATGGGAGCAATTA 1200
QY 1201 CTTGGCATGGTGGCAAGATAGAGACGTGTACTCGAGGGTCCCATTTTGGGCCAAAAT 1260
Db 1201 CTTGGCATGGTGGCAAGATAGAGACGTGTACTCGAGGGTCCCATTTTGGGCCAAAAT 1260
QY 1261 CTTCAACACAGATGGAACATTTTCAACCGTCTCTCTTATGGGCGGCTTTTGGACTCAAGAAC 1320
Db 1261 CTTCAACACAGATGGAACATTTTCAACCGTCTCTCTTATGGGCGGCTTTTGGACTCAAGAAC 1320
```

```
QY 1321 CGGCTCTCAGATCTCTCATCAAAAAACACGCTGTGTCTCGGAATCTCTCCGCGGAGTTT 1380
Db 1321 CGGCTCTCAGATCTCTCATCAAAAAACACGCTGTGTCTCGGAATCTCTCCGCGGAGTTT 1380
QY 1381 TCAGCTCAAAAAGTTTGTCTTCAATTCATCACCCAATACTCCACAGGACAAAGTGTGGAA 1440
Db 1381 TCAGCTCAAAAAGTTTGTCTTCAATTCATCACCCAATACTCCACAGGACAAAGTGTGGAA 1440
QY 1441 ATTGAATGGGAGCTGCGAGAAAAGAAAACAGCAAGCGCTGGAAATCCCGAAAGTGCAGTACACA 1500
Db 1441 ATTGAATGGGAGCTGCGAGAAAAGAAAACAGCAAGCGCTGGAAATCCCGAAAGTGCAGTACACA 1500
QY 1501 TCCAAATATGCAAAATCTGCCAACTGTTTACTGTGGACAAACAATGGACTTTTATCT 1560
Db 1501 TCCAAATATGCAAAATCTGCCAACTGTTTACTGTGGACAAACAATGGACTTTTATCT 1560
QY 1561 GAGCCTCCGCCCATTTGGCACCCCGTTTACCTTACCCTGCCCTGTA 1604
Db 1561 GAGCCTCCGCCCATTTGGCACCCCGTTTACCTTACCCTGCCCTGTA 1604
```

## RESULT 2

```
US-09-807-802A-14
; Sequence 14, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNYPN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1800
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1797)
; OTHER INFORMATION:
US-09-807-802A-14
```

```
Query Match 100.0%; Score 1604; DB 3; Length 1800;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGGCTTCAGGCGTGGCGCACCAATGCAGACAATAACGAAGGCGCCGACGGAGTGGGT 60
Db 196 ATGGCTTCAGGCGTGGCGCACCAATGCAGACAATAACGAAGGCGCCGACGGAGTGGGT 255
QY 61 AATGCTCAGGAAATTTGGCATTTCCATGCTGGCGGACAGAGTCAATCACCACC 120
Db 256 AATGCTCAGGAAATTTGGCATTTCCATGCTGGCGGACAGAGTCAATCACCACC 315
QY 121 AGCACCGGACCTGGGCGCTTGGCCACCTACATAACCACTCTTACAGCAAAATCTCCACT 180
Db 316 AGCACCGGACCTGGGCGCTTGGCCACCTACATAACCACTCTTACAGCAAAATCTCCACT 375
QY 181 GCTTCAACGGGGCGGACGAAACGACAAACCACTACTTTCGGCTACAGACCCCTCTGGGGGTAT 240
Db 376 GCTTCAACGGGGCGGACGAAACGACAAACCACTACTTTCGGCTACAGACCCCTCTGGGGGTAT 435
QY 241 TTTGATTTCAACAGATTTCACTGCGCACTTTTCCACCGTGTGAGCTGGCAGGACTCATCAAC 300
Db 436 TTTGATTTCAACAGATTTCCACTGGCACTTTTCCACCGTGTGAGCTGGCAGGACTCATCAAC 495
```

QY 301 AACAAATGGGGATTCCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360  
DB 496 AACAAATGGGGATTCCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 355  
QY 361 GAGGTACAGAGAAATGATGGCGTCAACAACATCGCTTAATACCTTTACAGACGGTTCAA 420  
DB 556 GAGGTACAGAGAAATGATGGCGTCAACAACATCGCTTAATACCTTTACAGACGGTTCAA 615  
QY 421 GTCTTCTCGGACTCGGAGTACCAAGCTTCCGTACGTCTCGGCTCTCGGCACCAAGGCTGC 480  
DB 616 GTCTTCTCGGACTCGGAGTACCAAGCTTCCGTACGTCTCGGCTCTCGGCACCAAGGCTGC 675  
QY 481 CTCCTCTCGGTTCCCGCGGAGCTGTTTCATGATTCGGAATACGGCTACCTGACGCTCAAC 540  
DB 676 CTCCTCTCGGTTCCCGCGGAGCTGTTTCATGATTCGGAATACGGCTACCTGACGCTCAAC 735  
QY 541 AATGGCAGCCAGCCGCTGGAGCTTCACTCTTTACTGCTGGAATATTTCCCTTCTCAG 600  
DB 736 AATGGCAGCCAGCCGCTGGAGCTTCACTCTTTACTGCTGGAATATTTCCCTTCTCAG 795  
QY 601 ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660  
DB 796 ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 855  
QY 661 AGCAGCTACGCGCACAGCCGCTGAGACCTGAGACCGCTGATGAATCCTCTCATCGACCAATAC 720  
DB 856 AGCAGCTACGCGCACAGCCGCTGAGACCGCTGATGAATCCTCTCATCGACCAATAC 915  
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTTACCAAGGACTTGTCTG 780  
DB 916 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTTACCAAGGACTTGTCTG 975  
QY 781 TTTAGCCGTGGGTCTCCAGCTGCGATGCTGTTCAGCCCAAAACTGGCTACCTGGACCC 840  
DB 976 TTTAGCCGTGGGTCTCCAGCTGCGATGCTGTTCAGCCCAAAACTGGCTACCTGGACCC 1035  
QY 841 TGTATTCCGCGAGCGCGCTTCTTAAACAAACACAGACAAACACAGCAATTTTACC 900  
DB 1036 TGTATTCCGCGAGCGCGCTTCTTAAACAAACACAGACAAACACAGCAATTTTACC 1095  
QY 901 TGGACTGGTGTCTTCAAAATTAACCTCAATGGGGGTGAATCGATCATCAACCTTGGCACT 960  
DB 1096 TGGACTGGTGTCTTCAAAATTAACCTCAATGGGGGTGAATCGATCATCAACCTTGGCACT 1155  
QY 961 GCTATGSCCTCACAAAAGACGACGAGCAAGTCTTCCCATGAGCGGTGTCATGATT 1020  
DB 1156 GCTATGSCCTCACAAAAGACGACGAGCAAGTCTTCCCATGAGCGGTGTCATGATT 1215  
QY 1021 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATTTGGAACAATGTTCATGATTACAGAC 1080  
DB 1216 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATTTGGAACAATGTTCATGATTACAGAC 1275  
QY 1081 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTGGGACCGTGGCGATC 1140  
DB 1276 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTGGGACCGTGGCGATC 1335  
QY 1141 AATTTCCAGACGACGACGACGACCCCTGCGACCGGAGATGTCATGCTATGGGACATTA 1200  
DB 1336 AATTTCCAGACGACGACGACGACCCCTGCGACCGGAGATGTCATGCTATGGGACATTA 1395  
QY 1201 CCTGGCATGTGTGGCAAGATAGAGACGTGTACTCTGACGGGTCCCAATTTGGGCCAAAATT 1260  
DB 1396 CCTGGCATGTGTGGCAAGATAGAGACGTGTACTCTGACGGGTCCCAATTTGGGCCAAAATT 1455  
QY 1261 CCTCACACAGATGGACATTTTCAACCGCTCTCTCTTATAGGCGGCTTTTGGACTCAAGAAC 1320  
DB 1456 CCTCACACAGATGGACATTTTCAACCGCTCTCTCTTATAGGCGGCTTTTGGACTCAAGAAC 1515  
QY 1321 CCGCCTCTCAGATCTCATCAAAAACAGCCTGTTCTCGGAATCTCTCGGCGGAGTTT 1380  
DB 1516 CCGCCTCTCAGATCTCTCATCAAAAACAGCCTGTTCTCGGAATCTCTCGGCGGAGTTT 1575  
QY 1381 TCAGCTACAAAGTTTGTCTTCAATCATCAACCAATFACTCCACAGACAAAGTGAGTGTGAA 1440

DB 1576 TCAGCTACAAAGTTTGTCTTCAATCATCCCAATACTCCACAGGACAAGTGAGTGTGAA 1635  
QY 1441 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500  
DB 1636 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1695  
QY 1501 TCAATATTGCAAAATCTGCCAAAGTGTGATTTTACTGTGGACAAACAATGGACTTTTACT 1560  
DB 1696 TCAATATTGCAAAATCTGCCAAAGTGTGATTTTACTGTGGACAAACAATGGACTTTTACT 1755  
QY 1561 GAGCCTCGCCCATTTGGCACCCGTTACCTTACCCTCCCTGTA 1604  
DB 1756 GAGCCTCGCCCATTTGGCACCCGTTACCTTACCCTCCCTGTA 1799

## RESULT 3

US-09-807-802A-12  
; Sequence 12, Application US/09807802A  
; Patent No. 6759237  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,  
; FILE REFERENCE: GNPVN.031USA  
; CURRENT APPLICATION NUMBER: US/09/807,802A  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/107,114  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: PCT/US99/25694  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 12  
; LENGTH: 2211  
; TYPE: DNA  
; ORGANISM: AAV-1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2208)  
; OTHER INFORMATION:  
US-09-807-802A-12

Query Match 100.0%; Score 1604; DB 3; Length 2211;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTTCAGGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCCCGCAGGAGTGGGT 60  
DB 607 ATGGCTTCAGGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCCCGCAGGAGTGGGT 666  
QY 61 AATGCCTCAGGAAATTTGGCATTTCCGATTTCCACATGGCTGGCGCAGAGTTCATCACACC 120  
DB 667 AATGCCTCAGGAAATTTGGCATTTCCGATTTCCACATGGCTGGCGCAGAGTTCATCACACC 726  
QY 121 AGCACCCGACCTGGGSCCTTGGCCACCTCAATAAACACCTCTAAGCAAAATCTCCAGT 180  
DB 727 AGCACCCGACCTGGGSCCTTGGCCACCTCAATAAACACCTCTAAGCAAAATCTCCAGT 786  
QY 181 GCTTCAACGGGGGCCAGCAACGACCACTACTCTCGGCTACAGCACCCCTCGGGGTAT 240  
DB 787 GCTTCAACGGGGGCCAGCAACGACCACTACTCTCGGCTACAGCACCCCTCGGGGTAT 846  
QY 241 TTTGATTTCAACAGATTTCCACTGCGCATTTTACAGAGTACTGGCAGGAGTTCATCAAC 300  
DB 847 TTTGATTTCAACAGATTTCCACTGCGCATTTTACAGAGTACTGGCAGGAGTTCATCAAC 906  
QY 301 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360  
DB 907 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 966  
QY 361 GAGGTACAGCAAGTAATGATGGCGTGCACAAACCATCGCTAATAACCTTACAGCAGCGTTCAA 420

Db 967 GAGGTACGACGAATGATGGCGTCACACCATCGCTAATACCTTACCGACGGTTCAA 1026  
QY 421 GTCTTCTCGGACTCGGAGTACAGGCTTCGTAAGTCTCGGCTTCGGGACCGAGGCTGC 480  
Db 1027 GTCTTCTCGGACTCGGAGTACAGGCTTCGTAAGTCTCGGCTTCGGGACCGAGGCTGC 1086  
QY 481 CTCCTCTCGGTTCCCGGCGGAGTGTTCATGATTCGGCAATACGGCTACGCTCAAC 540  
Db 1087 CTCCTCTCGGTTCCCGGCGGAGTGTTCATGATTCGGCAATACGGCTACGCTCAAC 1146  
QY 541 AATGCGACGAAGCGGTGGGAGCTTCATCCTTTTACTGCTGGAATATTTCCCTTCTCAG 600  
Db 1147 AATGCGACGAAGCGGTGGGAGCTTCATCCTTTTACTGCTGGAATATTTCCCTTCTCAG 1206  
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACTTTTGAGGAAGTGCCTTTCCAC 660  
Db 1207 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACTTTTGAGGAAGTGCCTTTCCAC 1266  
QY 661 AGCAGCTACGCGACAGCAGGAGCTGACCGGCTGATGATCCCTCTCATCGACCAATAC 720  
Db 1267 AGCAGCTACGCGACAGCAGGAGCTGACCGGCTGATGATCCCTCTCATCGACCAATAC 1326  
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTGCAAAACAAAGGACTTGTG 780  
Db 1327 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTGCAAAACAAAGGACTTGTG 1386  
QY 781 TTTAGCCTGGTGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACTCGAGCC 840  
Db 1387 TTTAGCCTGGTGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACTCGAGCC 1446  
QY 841 TGTTATCGCAGCAGGCGGTCTTAAACAAAAAAGAGCAACAAACAGCAATTTTACC 900  
Db 1447 TGTTATCGCAGCAGGCGGTCTTAAACAAAAAAGAGCAACAAACAGCAATTTTACC 1506  
QY 901 TGGACTGGTGTCTTAAATATATACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 960  
Db 1507 TGGACTGGTGTCTTAAATATATACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 1566  
QY 961 GCTATGGCTCTCACAAAGACAGCAAGTGTCTTTCCATGAGCGGTGTCAATGATT 1020  
Db 1567 GCTATGGCTCTCACAAAGACAGCAAGTGTCTTTCCATGAGCGGTGTCAATGATT 1626  
QY 1021 TTTGGAAGAGAGCGCGGAGCTTCAACATCTGATTTGGCAATGTCTATGATTCAGAC 1080  
Db 1627 TTTGGAAGAGAGCGCGGAGCTTCAACATCTGATTTGGCAATGTCTATGATTCAGAC 1686  
QY 1081 GAAGAGGAATTTAAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140  
Db 1687 GAAGAGGAATTTAAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1746  
QY 1141 AATTTCCAGAGCAGCAGACACCTCGACCGGAGATGTGCATGTATGGAGCATTA 1200  
Db 1747 AATTTCCAGAGCAGCAGACACCTCGACCGGAGATGTGCATGTATGGAGCATTA 1806  
QY 1201 CTGGCATGGTGGCAAGATAGAGACGTGTACTGCGAGGTGCCATTTTGGGCCAAAATT 1260  
Db 1807 CTGGCATGGTGGCAAGATAGAGACGTGTACTGCGAGGTGCCATTTTGGGCCAAAATT 1866  
QY 1261 CTTCAACAGATGGACACTTTTACCCGTCTCTCTTATGGCGGCTTTTGGACTCAAGAAC 1320  
Db 1867 CTTCAACAGATGGACACTTTTACCCGTCTCTCTTATGGCGGCTTTTGGACTCAAGAAC 1926  
QY 1321 CGCCTCTCTCAGATCTCTCAAAAAACAGCCTGTTCCTGCGAATCCTCCGCGGAGTTT 1380  
Db 1927 CGCCTCTCTCAGATCTCTCAAAAAACAGCCTGTTCCTGCGAATCCTCCGCGGAGTTT 1986  
QY 1381 TCAGCTACAAAGTTTCTTCAATTCACCCAACTACTCCACAGGACAAAGTGTGGAA 1440  
Db 1987 TCAGCTACAAAGTTTCTTCAATTCACCCAACTACTCCACAGGACAAAGTGTGGAA 2046  
QY 1441 ATTGAATGGAGCTGCAGAAAGAAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500

Db 2047 ATTGAATGGAGCTGCAGAAAGAAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 2106  
QY 1501 TCCAATTATGCAAAATCTGCCAAGCTTGAATTTTACTGTGGACAAATGGACTTTATACT 1560  
Db 2107 TCCAATTATGCAAAATCTGCCAAGCTTGAATTTTACTGTGGACAAATGGACTTTATACT 2166  
QY 1561 GAGCCTCGCCCCATTTGGCACCCGTTACCTTACCCCGTCCCTGTGA 1604  
Db 2167 GAGCCTCGCCCCATTTGGCACCCGTTACCTTACCCCGTCCCTGTGA 2210  
  
RESULT 4  
US-09-807-802A-1  
; Sequence 1, Application US/09807802A  
; Patent No. 6759237  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,  
; TITLE OF INVENTION: Vectors and Host Cells Containing Same  
; FILE REFERENCE: GNPVN.031USA  
; CURRENT APPLICATION NUMBER: US/09/807,802A  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/107,114  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: PCT/US99/25694  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 4718  
; TYPE: DNA  
; ORGANISM: AAV-1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (335)..(2206)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2223)..(4430)  
; OTHER INFORMATION:  
US-09-807-802A-1  
  
Query Match 100.0%; Score 1604; DB 3; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAAATAACGAAGGCGCGCAGGAGTGGGT 60  
Db 2829 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAAATAACGAAGGCGCGCAGGAGTGGGT 2888  
  
QY 61 AATGCTCAGGAATTTGGCATTTGCCATTTCCATGGCTGGGCGCAGAGTCAATCACCAC 120  
Db 2889 AATGCTCAGGAATTTGGCATTTGCCATTTCCATGGCTGGGCGCAGAGTCAATCACCAC 2948  
  
QY 121 AGCACCCGCACTGGGCGCTTGGCCACCTACATAACCACTTCCGCTCAGCACCCCTTGGGGGTAT 180  
Db 2949 AGCACCCGCACTGGGCGCTTGGCCACCTACATAACCACTTCCGCTCAGCACCCCTTGGGGGTAT 3008  
  
QY 181 GCTTCAACGGGGGCGCAGCAACGACAAACCACTTCTCGCTCAGCACCCCTTGGGGGTAT 240  
Db 3009 GCTTCAACGGGGGCGCAGCAACGACAAACCACTTCTCGCTCAGCACCCCTTGGGGGTAT 3068  
  
QY 241 TTTGATTTCAACAGATTCATCGCCATTTTTCACCGTGTGAGCTGGCAGGACTCATCAAC 300  
Db 3069 TTTGATTTCAACAGATTCATCGCCATTTTTCACCGTGTGAGCTGGCAGGACTCATCAAC 3128  
  
QY 301 AACAAATGGGATTCGGGCGCCAGAGACTCAACTTTCAACTCTTCAACATCAAGTCAAG 360  
Db 3129 AACAAATGGGATTCGGGCGCCAGAGACTCAACTTTCAACTCTTCAACATCAAGTCAAG 3188  
  
QY 361 GAGGTACAGACGAATGATGGGTCAACACCATCGCTTAATTAACCTTACCAGCACGGTTCAA 420

Db 3189 GAGGTCAACGAAATGATGGCGTCAACAACATCGCTAAATAACCTTACCGACGACGGTTCAA 3248  
QY 421 GTCTTCTCGGACTCGGAGTACAGACTTTCGGTACGTCTCGGGTCTCGCGACACGAGGCTGC 480  
Db 3249 GTCTTCTCGGACTCGGAGTACAGACTTTCGGTACGTCTCGGGTCTCGCGACACGAGGCTGC 3308  
QY 481 CTCCCTCCGTTCGCCGCGACGTTGTCATGATTCGCAATACGCTACCTGACGCTCAAC 540  
Db 3309 CTCCCTCCGTTCGCCGCGACGTTGTCATGATTCGCAATACGCTACCTGACGCTCAAC 3368  
QY 541 AATGGCAGCAAGCCGTCGGAGCGTTTCATCTCTTTACTGCTGGAATATTTCCCTTCTCAG 600  
Db 3369 AATGGCAGCAAGCCGTCGGAGCGTTTCATCTCTTTACTGCTGGAATATTTCCCTTCTCAG 3428  
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660  
Db 3429 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 3488  
QY 661 AGCAGCTACGCGCACAGCGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720  
Db 3489 AGCAGCTACGCGCACAGCGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 3548  
QY 721 CTGTATTACTGTAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAACAGAGACTTGCCTG 780  
Db 3549 CTGTATTACTGTAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAACAGAGACTTGCCTG 3608  
QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAACCTGGCTACCTGGACCC 840  
Db 3609 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAACCTGGCTACCTGGACCC 3668  
QY 841 TGTATTCCGACAGCGCGGTTCTTAAACAAAACAGACAAACAAACAGCAATTTTACC 900  
Db 3669 TGTATTCCGACAGCGCGGTTCTTAAACAAAACAGACAAACAAACAGCAATTTTACC 3728  
QY 901 TGGACTGTGTCTCAAAATATACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 960  
Db 3729 TGGACTGTGTCTCAAAATATACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 3788  
QY 961 GCTATGGCTCTACAACAGACGACGAAAGTCTTCCATGAGCGGTGTCATGATT 1020  
Db 3789 GCTATGGCTCTACAACAGACGACGAAAGTCTTCCATGAGCGGTGTCATGATT 3848  
QY 1021 TTTGAAAAAGAGAGCGCCGAGCTTCAACACACTGCAATGGACAATGTATGATTACAGAC 1080  
Db 3849 TTTGAAAAAGAGAGCGCCGAGCTTCAACACACTGCAATGGACAATGTATGATTACAGAC 3908  
QY 1081 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTGGGACCGTGGCAGTC 1140  
Db 3909 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTGGGACCGTGGCAGTC 3968  
QY 1141 AATTTCCAGACGACGACAGACCTTGGACCGGAGATGTGATGCTATGGGAGCATTA 1200  
Db 3969 AATTTCCAGACGACGACAGACCTTGGACCGGAGATGTGATGCTATGGGAGCATTA 4028  
QY 1201 CTTGGCATGTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTTGGGCCAAAAAT 1260  
Db 4029 CTTGGCATGTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTTGGGCCAAAAAT 4088  
QY 1261 CCTCACACAGATGGACACTTTTACCCGTCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 1320  
Db 4089 CCTCACACAGATGGACACTTTTACCCGTCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 4148  
QY 1321 CCGCTCTCTCAGATCCTCATCAAAAACAGCCCTGTTCTCCGAATCCTCCGGCGGAGTTT 1380  
Db 4149 CCGCTCTCTCAGATCCTCATCAAAAACAGCCCTGTTCTCCGAATCCTCCGGCGGAGTTT 4208  
QY 1381 TCAGCTACAAGTTTGTCTTCAATTCATCACCCAAATCTCCAAGGACAAGTGAAGTGGAA 1440  
Db 4209 TCAGCTACAAGTTTGTCTTCAATTCATCACCCAAATCTCCAAGGACAAGTGAAGTGGAA 4268  
QY 1441 ATTGAATGGAGCTGCAGAAAGAAAAACAGAACGCTGGAAATCCCGAAGTGCAAGTACACA 1500  
Db 4269 ATTGAATGGAGCTGCAGAAAGAAAAACAGAACGCTGGAAATCCCGAAGTGCAAGTACACA 4328

RESULT 5

US-10-216-870-11  
; Sequence 11, Application US/10216870  
; Patent No. 6723551  
; GENERAL INFORMATION:  
; APPLICANT: KOTIN, ROBERT M  
; APPLICANT: URABE, MASASHI  
; APPLICANT: DING, CHUAN-TIAN  
; TITLE OF INVENTION: PRODUCTION OF ADENO-ASSOCIATED VIRUS IN INSECT CELLS  
; FILE REFERENCE: 402133  
; CURRENT APPLICATION NUMBER: US/10/216,870  
; CURRENT FILING DATE: 2002-08-13  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11  
; LENGTH: 7447  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-216-870-11

Query Match 100.0%; Score 1604; DB 3; Length 7447;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTTCAGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCCGACGAGTGGGT 60  
Db 5227 ATGGCTTCAGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCCGACGAGTGGGT 5286  
QY 61 AATGCTCAGGAAATTTGGCATTCGCAATGCCATTCACATGGCTGGGCGACAGAGTCAACACC 120  
Db 5287 AATGCTCAGGAAATTTGGCATTCGCAATGCCATTCACATGGCTGGGCGACAGAGTCAACACC 5346  
QY 121 AGCACCCGACCTGGGCTTGGCCACCTACAAATAACCACTCTCAAGCAAAATCTCCAGT 180  
Db 5347 AGCACCCGACCTGGGCTTGGCCACCTACAAATAACCACTCTCAAGCAAAATCTCCAGT 5406  
QY 181 GCTTCAACGGGGCCAGCAACCACTACTCTTGGCTACAGCAACCCCTGGGGGTAT 240  
Db 5407 GCTTCAACGGGGCCAGCAACCACTACTCTTGGCTACAGCAACCCCTGGGGGTAT 5466  
QY 241 TTTGATTTCAACAGATTCCACCTGGCACTTTCACACGCTGACTGGCAGCGACTCATCAAC 300  
Db 5467 TTTGATTTCAACAGATTCCACCTGGCACTTTCACACGCTGACTGGCAGCGACTCATCAAC 5526  
QY 301 AACAAATTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAACTCAAG 360  
Db 5527 AACAAATTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAACTCAAG 5586  
QY 361 GAGGTCAACGAATGATGGGTCACAAACATCGCTTAATAACCTTACAGACGAGTTCAG 420  
Db 5587 GAGGTCAACGAATGATGGGTCACAAACATCGCTTAATAACCTTACAGACGAGTTCAG 5646  
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCCGTACGCTCTCGGTCTCGGCTCTCGGCAACGAGGCTGC 480  
Db 5647 GTCTTCTCGGACTCGGAGTACAGCTTCCGTACGCTCTCGGTCTCGGCTCTCGGCAACGAGGCTGC 5706  
QY 481 CTCCCTCCGTTCGCCGCGACGTTGTCATGATTCCGCAATACGGCTACCTGACGCTCAAC 540  
Db 5707 CTCCCTCCGTTCGCCGCGACGTTGTCATGATTCCGCAATACGGCTACCTGACGCTCAAC 5766  
QY 541 AATGGCAGCAAGCCGTCGGAGCGTTTCATCTTTACTGCTGGAATATTTTCCCTTCTCAG 600

```
Db 5767 AATGCGAGCAAGCGTGGAGGTTTCATCTTTTACTGCTGGATATTTCCCTTCTAG 5826
QY 601 ATGCTGAGAACGGGCAACATTTTACCTTACGCTACACCTTTGAGGAAGTGCTTTCCAC 660
Db 5827 ATGCTGAGAACGGGCAACATTTTACCTTACGCTACACCTTTGAGGAAGTGCTTTCCAC 5886
QY 661 AGCAGTACGGCACAGCAGCAGCCTGGACCGGCTGATGAATCCTCTCATGACCAATAC 720
Db 5887 AGCAGTACGGCACAGCAGCAGCCTGGACCGGCTGATGAATCCTCTCATGACCAATAC 5946
QY 721 CTGTATTACTCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAAGGACTTGCTG 780
Db 5947 CTGTATTACTCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAAGGACTTGCTG 6006
QY 781 TTTAGCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAAACCTGGCTACTCGGACCC 840
Db 6007 TTTAGCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAAACCTGGCTACTCGGACCC 6066
QY 841 TGTATTCCGCGAGCGCGGTCTTAAACAAAAAAGACAAACAAAGCAATTTTACC 900
Db 6067 TGTATTCCGCGAGCGCGGTCTTAAACAAAAAAGACAAACAAAGCAATTTTACC 6126
QY 901 TGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCTCGCACT 960
Db 6127 TGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCTCGCACT 6186
QY 961 GCTATGGCTCTCACAAAGACGACGAAGACAAGTTCTTTTCCATGAGCGGTGTCTATGATT 1020
Db 6187 GCTATGGCTCTCACAAAGACGACGAAGACAAGTTCTTTTCCATGAGCGGTGTCTATGATT 6246
QY 1021 TTTGGAAAGAGAGCGCGGAGCTTCAAACTGCAATTTGGACAAATGTCTATGATTACAGAC 1080
Db 6247 TTTGGAAAGAGAGCGCGGAGCTTCAAACTGCAATTTGGACAAATGTCTATGATTACAGAC 6306
QY 1081 GAAGAGGAAATTAAGCCACTTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140
Db 6307 GAAGAGGAAATTAAGCCACTTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 6366
QY 1141 AATTTCCAGAGCAGCAGCAGACAGCCCTGCGACCGGAGATGTGCTATGAGGAGATT 1200
Db 6367 AATTTCCAGAGCAGCAGCAGACAGCCCTGCGACCGGAGATGTGCTATGAGGAGATT 6426
QY 1201 CCTGGCATGGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTTGGGCCAAAATT 1260
Db 6427 CCTGGCATGGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTTGGGCCAAAATT 6486
QY 1261 CCTCACACAGATGGACACTTTTACCCGCTCTCTCTTTATGGCGGCTTTGGACCTCAAGAAC 1320
Db 6487 CCTCACACAGATGGACACTTTTACCCGCTCTCTCTTTATGGCGGCTTTGGACCTCAAGAAC 6546
QY 1321 CGGCTCTCTCAGATCTCATCAAAAAACAGCCTGTGTTCTTGGCAATCCTCCGCGGAGTTT 1380
Db 6547 CGGCTCTCTCAGATCTCATCAAAAAACAGCCTGTGTTCTTGGCAATCCTCCGCGGAGTTT 6606
QY 1381 TCAGCTACAAAGTTTGCTTTCATTCACCCATCTCCACAGGACAAAGTGTGGAA 1440
Db 6607 TCAGCTACAAAGTTTGCTTTCATTCACCCATCTCCACAGGACAAAGTGTGGAA 6666
QY 1441 ATTGAATGGAGCTGCAGAAAGAAAACAGACGCTGGAATCCCGAAGTGCAATACACA 1500
Db 6667 ATTGAATGGAGCTGCAGAAAGAAAACAGACGCTGGAATCCCGAAGTGCAATACACA 6726
QY 1501 TCAATATTGCAAAATCTGCCAAAGTTGATTTTATCTGTGGACAAACAAATGAGCTTTATCT 1560
Db 6727 TCAATATTGCAAAATCTGCCAAAGTTGATTTTATCTGTGGACAAACAAATGAGCTTTATCT 6786
QY 1561 GAGCTTGGCCCAATTTGGGACCGGTTACCTTACCGTCCCTGTGA 1604
Db 6787 GAGCTTGGCCCAATTTGGGACCGGTTACCTTACCGTCCCTGTGA 6830
```

```
US-09-807-802A-19
; Sequence 19, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,
; CURRENT APPLICATION NUMBER: US/09/807,802A
; FILE REFERENCE: GNVN.031USA
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 4683
; TYPE: DNA
; ORGANISM: AAV-6
US-09-807-802A-19
```

```
Query Match 94.3%; Score 1512.8; DB 3; Length 4683;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1547; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAAGCGCCGACGAGTGGGT 60
Db 2814 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAAGCGCCGACGAGTGGGT 2873
QY 61 AATGCTTCAGGAATTTGGCATTTGGCATTTCCATATGGCTTGGGGGACAGAGTCTATCACC 120
Db 2874 AATGCTTCAGGAATTTGGCATTTGGCATTTCCATATGGCTTGGGGGACAGAGTCTATCACC 2933
QY 121 AGCACCCGACCTGGGCGCTTCCCACTTACCAATTAACCACTCTCAACGCAAAATCTCCAGT 180
Db 2934 AGCACCCGAAATGGGCGCTTCCCACTTACCAATTAACCACTCTCAACGCAAAATCTCCAGT 2993
QY 181 GCTTCAACGGGGGCGCAGCAACGACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT 240
Db 2994 GCTTCAACGGGGGCGCAGCAACGACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT 3053
QY 241 TTTGATTTCAACAGATTTCCACTTGGCACTTTTTCACCACTGACTGGCAGCGACTCATCAAC 300
Db 3054 TTTGATTTCAACAGATTTCCACTTGGCACTTTTTCACCACTGACTGGCAGCGACTCATCAAC 3113
QY 301 AACAAATGGGGAATCCGCGCCCAAGAGACTCAACTTCAAACCTCTTCAAACATCCAAGTCAAG 360
Db 3114 AACAAATGGGGAATCCGCGCCCAAGAGACTCAACTTCAAACCTCTTCAAACATCCAAGTCAAG 3173
QY 361 GAGGTACGACGAAATGATGGGTGTCACAACTATCGCTAATAAACCCTTACCAGCACGGTTCAA 420
Db 3174 GAGGTACGACGAAATGATGGGTGTCACAACTATCGCTAATAAACCCTTACCAGCACGGTTCAA 3233
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGCTCTCGGCTCTGGCAGCAGCGGCTGC 480
Db 3234 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGCTCTCGGCTCTGGCAGCAGCGGCTGC 3293
QY 481 CTCCCTCTCGTTCGCGCGGAGCTGTTCATGATTTCCGCAATACGGCTACTCGACGCTCAAC 540
Db 3294 CTCCCTCTCGTTCGCGCGGAGCTGTTCATGATTTCCGCAATACGGCTACTCGACGCTCAAC 3353
QY 541 AATGCGACCCCAAGCGGTGGGACGTTTATCTCTTTTACTGCTGGAAATATTTCCCTTTCTCAG 600
Db 3354 AATGCGACCCCAAGCGGTGGGACGTTTATCTCTTTTACTGCTGGAAATATTTCCCTTTCTCAG 3413
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACTTTTGAGGAAGTGGCTTTCCAC 660
Db 3414 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACTTTTGAGGAAGTGGCTTTCCAC 3473
QY 661 AGCAGCTACGGCGCACAGCGCAGAGCTCGGCTGATGAATCCCTCTCATCGACCAATAC 720
```



```
Db 3474 AGCAGTACCGCACAGCAGCGAGCCTGGACCGGCTGATGAATCTCTCTCATCGACGAGTAC 3533
Qy 721 CTGTATTACTGAACAGAACTCAAAATCAGTCGGAGTGCACAAACAAAGACTTGTCTG 780
Db 3534 CTGTATTACTGACAGAACTCACAATCAGTCGGAGTGCACAAACAAAGACTTGTCTG 3593
Qy 781 TTATGCGGTGCTCCAGCTGGCATGTCTGTTTCCAGCCCAAAACCTGGCTACCTGACCC 840
Db 3594 TTATGCGGTGCTCCAGCTGGCATGTCTGTTTCCAGCCCAAAACCTGGCTACCTGACCC 3653
Qy 841 TGTATTCCGACGCGCGCTTCTAATAACAAACAAACAGACAAACAAACGAAATTTTACC 900
Db 3654 TGTATTCCGACGCGCGCTTCTAATAACAAACAAACAGACAAACAAACGAAATTTTACC 3713
Qy 901 TGGACTGTGCTTCAAAATATAACCTCAATGGCGGTGAATCATCATCAACCTGGCACT 960
Db 3714 TGGACTGTGCTTCAAAATATAACCTCAATGGCGGTGAATCATCATCAACCTGGCACT 3773
Qy 961 GCTATGGCTCTCACAAAGACGACGAAGAACAAGTTCTTCCATGAGCGGTGTCAATGAT 1020
Db 3774 GCTATGGCTCTCACAAAGACGACGAAGAACAAGTTCTTCCATGAGCGGTGTCAATGAT 3833
Qy 1021 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGACAAATGTCAATGATACAGAC 1080
Db 3834 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGACAAATGTCAATGATACAGAC 3893
Qy 1081 GAAGAGAAATTAAGCCCACTAACCTGTGGCCGACGAAAGATTTGGGACCGTGGCAGTC 1140
Db 3894 GAAGAGAAATTAAGCCCACTAACCTGTGGCCGACGAAAGATTTGGGACCGTGGCAGTC 3953
Qy 1141 AATTTCCAGAGCAGCAGCAGACAGACCCCTGGCAGCGAGATGTGATGCTATGGGAGCATTA 1200
Db 3954 AATTTCCAGAGCAGCAGCAGACAGACCCCTGGCAGCGAGATGTGATGCTATGGGAGCATTA 4013
Qy 1201 CTTGGCATGGTGGCAAGATAGAGACGCTGACCTGCGAGGTCCTCAATTTGGGCAAAAT 1260
Db 4014 CTTGGCATGGTGGCAAGATAGAGACGCTGACCTGCGAGGTCCTCAATTTGGGCAAAAT 4073
Qy 1261 CTTCCACAGATGGACACTTTTCAACCGCTCTCTCTTATGGGCGCTTTGGACTCAAGAAC 1320
Db 4074 CTTCCACAGATGGACACTTTTCAACCGCTCTCTCTTATGGGCGCTTTGGACTCAAGAAC 4133
Qy 1321 CGGCTCTCTCAGATCTCTCAAAAAACAGCGCTGTTCTTCCGAAATCCTCGGCGGAGTTT 1380
Db 4134 CGGCTCTCTCAGATCTCTCAAAAAACAGCGCTGTTCTTCCGAAATCCTCGGCGGAGTTT 4193
Qy 1381 TCAGCTACAAAGTTTGGCTTCAATTCATCACCCAAATATCTCCAAGGAAAGTGTGGAA 1440
Db 4194 TCGGCTACAAAGTTTGGCTTCAATTCATCACCCAGTATTCACAGGAAAGTGTGGAG 4253
Qy 1441 ATTGAATGGAGCTGCAGAAAGAAACAGCAGCGCTGGAAATCCCGAAGTGCAGTACACA 1500
Db 4254 ATTGAATGGAGCTGCAGAAAGAAACAGCAGCGCTGGAAATCCCGAAGTGCAGTACACA 4313
Qy 1501 TCCAAATATGCAAAATCTGCAACAGTTGATTTTACTGTGGACAAACAAATGACTTTTACT 1560
Db 4314 TCTAATATGCAAAATCTGCAACAGTTGATTTTACTGTGGACAAACAAATGACTTTTACT 4373
Qy 1561 GAGCCTCGCCCAATGGCACCCGTTACTTACCTTACCCGTCCTGTA 1604
Db 4374 GAGCCTCGCCCAATGGCACCCGTTACTTACCTTACCCGTCCTGTA 4417
```

RESULT 7

```
US-09-770-315-4
; Sequence 4, Application US/09770315
; Patent No. 6429001
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Recombinant AAV Packaging Systems
; FILE REFERENCE: 20263-501
; CURRENT APPLICATION NUMBER: US/09/770,315
; CURRENT FILING DATE: 2001-01-26
```

```
; PRIOR APPLICATION NUMBER: US 60/178,536
; PRIOR FILING DATE: 2000-01-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 4072
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: recombinant DNA
US-09-770-315-4

Query Match      64.0%; Score 1026.4; DB 3; Length 4072;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 1250; Conservative 0; Mismatches 351; Indels 3; Gaps 1;

Qy 1 ATGGCTTCAGGCGGTGGCGCAATGCGACCAATGCGACAAATAACGAGGCGCGACGAGTGGGT 60
Db 2090 ATGGCTTCAGGCGGTGGCGCAATGCGACCAATGCGACCAATGCGACGAGTGGGT 2149
Qy 61 AATGCTTCAGGAAATTTGGCATTTGCCATTTCCACATGATGGGCGACAGAGTCTATCCACC 120
Db 2150 AATTCCTCGGAAATTTGGCATTTGCCATTTCCACATGATGGGCGACAGAGTCTATCCACC 2209
Qy 121 AGCACCGCACCTGGGCTTGGCCACCTTACATAACACCTCTACAGCAAAATCTCCAGT 180
Db 2210 AGCACCGCAACTGGGCTTGGCCACCTTACATAACACACCTCTACAGCAAAATTTCCAGC 2269
Qy 181 GCTTCAACGGGGCGGCGACGACAAACACACTTCTCGGCTACAGCACCCCTGGGGGTAT 240
Db 2270 CAATCA---GGAGCTTCGAACGACCAATCACTATTGGCTACAGCACCCCTTGGGGGTAT 2326
Qy 241 TTGTGATTTCAACAGATTTCCACTTGGCCACTTTTCCACCACTGCTGAGTGGCAGCGACTCATCAAC 300
Db 2327 TTGTGATTTCAACAGATTTCCACTTGGCCACTTTTCCACCACTGCTGAGTGGCAGCGACTCATCAAC 2386
Qy 301 AACAAATGGGGATTCGGGCGGCAAGAGACTCAACTTCAAACTCTTCAAGTCAAGTCAAG 360
Db 2387 AACAACTGGGGATTCGGGCGGCAAGAGACTCAACTTCAAACTCTTCAAGTCAAGTCAAA 2446
Qy 361 GAGTTCACGACCAAGATGATGGCGTCAACACCTGCTATAAAGCTTACACGACGCTTCAA 420
Db 2447 GAGTTCACGACCAAGATGATGGCGTCAACACCTGCTATAAAGCTTACACGACGCTTCAA 2506
Qy 421 GTCTTCTCGGACTCGGAGTACCGAGTTCGCTACGCTCTCGGCTCTGCGCACCGAGCTGC 480
Db 2507 GTGTCTTCTCGGACTCGGAGTACCGAGTTCGCTACGCTCTCGGCTCGGCGCATCAAGATGC 2566
Qy 481 CTCTCTTCTCGGCTTCGGGCGGACGCTGTTCAATGATTCGCGCAATACGGCTACCTGAGCTCAAC 540
Db 2567 CTCTCTTCTCGGCTTCGGGCGGACGCTGTTCAATGATTCGCGCAATACGGCTACCTGAGCTCAAC 2626
Qy 541 AATGGCAGCAAGCGGTGGGAGCTTCACTTTCAGTACACCTTTTGGAGAAAGTGCCTTTCCAC 600
Db 2627 AATGGCAGCTACGCGAGTGGGAGCTTCACTTTCAGTACACCTTTTGGAGAAAGTGCCTTTCCAC 2686
Qy 601 ATGTCTGAGAAACGGGCAACAACTTTTACCTTTCAGTACACCTTTTGGAGAAAGTGCCTTTCCAC 660
Db 2687 ATGTCTGAGTACCGGAAACAACTTTTACCTTTCAGTACACCTTTTGGAGAAAGTGCCTTTCCAC 2746
Qy 661 AGCAGCTACGCGCACAGCGAGCTTGGACCGGCTGATGAAATCTCTCTATCGACCAATATC 720
Db 2747 AGCAGCTACGCTCACAGCGCAGAGTCTGGACCGCTCTCATGAATCTCTCATCGACCAATATC 2806
Qy 721 CTGTATTACTGAAACGAACTCAAAATCAGTCGGAAAGTGCACAAACAAAGACTTGTCTG 780
Db 2807 CTGTATTACTGAGCAGAAACAAACACTTCCAAAGTGGAAACCAACCGCATGCTCAAGGCTTCAG 2866
Qy 781 TTATGCGGTGGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAAACCTGGCTACTCTGACCC 840
Db 2867 TTATGCGGTGGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAGTGTCTGAGACTTCTGACCC 2926
Qy 841 TGTATTATCGGACGCGCGCTTTTCAAAACAAACAAACAGACAAACAAACAGCAATTTTACC 900
```



```
||||| 2927 TGTATCCGCGCAGCGAGTATCAAGACATCTCGGATATCAACACAGTGAATACTCG 2986
QY 901 TGGACTGGTGTCTTCAAAATATTAACCTCAATGGCGTGAATCCATCATCAACCCCTGGCACT 960
Db 2987 TGGACTGGAGCTTACCAGGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGCCG 3046
QY 961 GCTATGGCTCTACACAAAGACGACGAAAGCAAGTCTTTCCCATGAGCGGTGTCATGATT 1020
Db 3047 GCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTCCTCAGAGCGGGTTCCTCATC 3106
QY 1021 TTTTGAAGAGAGCGCGGAGCTTCAACACATGCTATGGCAATGTCATGATTACAGAC 1080
Db 3107 TTTTGAAGCAAGGCTCAGAGAAACAAATGGGACATTTGAAGAGTCATGATTACAGAC 3166
QY 1081 GAAGAGGAATTAAGGCCACTAACCTGTGGCCACCGAAAGATTGGGACCGTGGCAGTC 1140
Db 3167 GAAGAGGAATCAGGACAACTCCGTTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3226
QY 1141 AATTTCCAGAGCAGCAGCAGACCGCTGGACCGGAGATGTGCATGCTATGGGAGCATTA 1200
Db 3227 AACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAACACACAAGGCGTTCTT 3286
QY 1201 CCTGCAATGCTGTGCAAGATAGAGACGTGTACCTGCGAGGTCCTCATTTGGGCCAAAATT 1260
Db 3287 CCAGGATGGTCTGGCAGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 3346
QY 1261 CCTCACACAGATGGACACTTTTCAACCGCTCTCCTCTTATGGGCGGCTTTGGACTCAAGAAC 1320
Db 3347 CCACACCGGACGGACATTTTCAACCCCTCTCCCTCATGGGTGGATTGGACTTAAACAC 3406
QY 1321 CGGCTCTCAGATCTCTATCAAAACACGCTGTCTTCTGCGAATCTCTCGGCGGAGTTT 1380
Db 3407 CCTCTCCACAGATTTCTCATCAAGAACACCCCGGTACTCTGCGAATCTCTCGAACCTTTC 3466
QY 1381 TCAGCTACAAAGTTTGTCTTCAATTCATCCCAATATCTCCACAGGACAGTGAAGTGGAA 1440
Db 3467 AGTGGGCAAGTTTGTCTTCTTCTCATCACAGTACTCTCACGGACAGTCAAGCTGGAG 3526
QY 1441 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500
Db 3527 ATCGAGTGGAGCTGCAGAAAGAAACAGCAACGCTGGAATCCCGAATTCAGTACACT 3586
QY 1501 TCAATATTGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTATACT 1560
Db 3587 TCCAACTACAACAAGTCTGTTAATGTGGACTTTACTGTGGACACTAATGGCGTGTATTCA 3646
QY 1561 GAGCCTCGCCCATTTGGCACCCGTTTACCTTACCCGTCCTCTGTA 1604
Db 3647 GAGCCTCGCCCATTTGGCACCCAGATACCTGATCTGTAATCTGTA 3690
```

## RESULT 8

```
US-10-038-972A-12
/ Sequence 12, Application US/10038972A
/ Patent No. 6962815
/ GENERAL INFORMATION:
/ APPLICANT: J. Bartlett
/ TITLE OF INVENTION: AAV VECTORS AND METHODS
/ FILE REFERENCE: 28335/36996US
/ CURRENT APPLICATION NUMBER: US/10/038, 972A
/ PRIOR FILING DATE: 2002-01-04
/ PRIOR APPLICATION NUMBER: US 60/260,124
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 12
/ LENGTH: 4679
/ TYPE: DNA
/ ORGANISM: adeno-associated virus 2
US-10-038-972A-12
```

Query Match 64.0%; Score 1026.4; DB 3; Length 4679;

```
Best Local Similarity 77.9%; Pred. No. 0;
Matches 1250; Conservative 0; Mismatches 351; Indels 3; Gaps 1;

QY 1 ATGCGTTTCAGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCGCGACGAGTGGGT 60
Db 2809 ATGCGTACAGCAGCTGGCGCACCAATGGCAGACAATAACGAAGCGCGCGACGAGTGGGT 2868
QY 61 AATGCTTCAGAAATTTGGCATTTGGCATTTCCATATGGCTGGCGCAGAGTGTATCATCCACC 120
Db 2869 AATTCCTCGGAAATTTGGCATTTGGCATTTCCATATGGATGGCGCAGAGTGTATCATCCACC 2928
QY 121 AGCACCGCACCTCGGCTTCCGCCCTTACAAATAACACACCTCTACAGCAAAATCTCCAGT 180
Db 2929 AGCACCGCACTCGGCTTCCGCCCTTACAAATAACACACCTCTACAAACAATTTCCAGC 2988
QY 181 GCTTCAAGCGGCGGCGCAGCAACACCACTTCTCGGCTACAGCACCCCTCGGGGTAT 240
Db 2989 CAATCA---GGAGCTTCGAACGACAATCTACTTCTGGCTACAGCACCCCTTGGGGGTAT 3045
QY 241 TTTGATTTCAACAGATTTCCACTGCGCATTTTTCACACATGTGACTGGCAGCGACTCATCAAC 300
Db 3046 TTTGACTTTCAACAGATTTCCACTGCGCATTTTTCACACATGTGACTGGCAGCGACTCATCAAC 3105
QY 301 AACAAATGGGATTTCCGSCCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAAGTCAAG 360
Db 3106 AACAACTGGGATTTCCGACCCCAAGAGACTCAACTTTCAAGCTCTTTAAATTTCAAGTCAAA 3165
QY 361 GAGGTCAAGCAATGATGGCGTCAACACCATCTCTAATAACCTTTACAGCACCGTTTCAA 420
Db 3166 GAGGTCAAGCAATGATGGCGTCAACACCATCTCTAATAACCTTTACAGCACCGTTTCAA 3225
QY 421 GTCTTCTCGACTCGGAGTACAGGTTTCGTAAGTCTCTCGGCTCTGGCGCACAGGGGTGC 480
Db 3226 GTGTTTACTGACTCGGAGTACAGGTTTCGTAAGTCTCTCGGCTCTGGCGCACAGGGGTGC 3285
QY 481 CTCCCTCGTTCCTCGGCGGAGCTGTATGATTCGGAATACGGTACTCTCGAGCTCAAC 540
Db 3286 CTCCCGCGTTCCTCGGCGGAGCTGTATGATTCGGAATACGGTACTCTCGAGCTCAAC 3345
QY 541 AATGCGCACCAAGCGGTGGGAGCTTCACTCTTTTACTGCTCGGAATATTTCCCTTCTCAG 600
Db 3346 AACGGAGTCAAGGAGTAGGAGCTCTTTCACTTTTACTGCTCGGAGTACTTTCTTCTCAG 3405
QY 601 ATGCTGAGAACGGGCAACAACCTTTTACTTCACTTCACTTTGAGGAAAGTGTCTTTCAC 660
Db 3406 ATGCTGCGTACCGGAAACAACCTTTTACTTCACTTCACTTTTGGAGGAGCTTCTTTCAC 3465
QY 661 AGCAGCTACGCGCACAGCGCTGACCGGCTGATGAATCTCTCTCATCGACCAATAC 720
Db 3466 AGCAGCTACGCTACAGCGCAGGCTGAGCCGCTCTCATGAATCTCTCTCATCGACCGTAC 3525
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGGCCCAAAACAAGGACTTGTCTG 780
Db 3526 CTGTATTACTTTGAGCAGAAACAACAACCTCAAGTGGAAACCAACAGCAGTCAAGGCTTCA 3585
QY 781 TTTAGCCGTGGGTCTCCAGCTGGGATGTCTGTTCAGCCCCAAAACCTGGCTACTCGGACCC 840
Db 3586 TTTTCTCAGGCGCGGAGCGAGTGACATTCGGGACCACTAGGAACTGGCTCTCGGACCC 3645
QY 841 TGTTATCGCAGCAGCGGCTTTCTTAAACAACAAACAGACAAACAACAGCAATTTTACC 900
Db 3646 TGTATCCCGCAGCAGCGAGTATCAAGACATCTCGGGATTAACAACAAGTGAATACTCG 3705
QY 901 TGGACTGGTGTCTTCAAAATATAAACCCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 960
Db 3706 TGGACTGGAGCTTACCAGTACCACCTCAATGGCAGAGACTCTCTGTGTAATCCGGGCCG 3765
QY 961 GCTATGGCTCTACACAAAGACGACGAAAGTCTTTTCCCATGTAGCGGGTGTCTATGATT 1020
Db 3766 GCCATGGCAAGCCAACAAGACGATGAAGAAAGTTTTCCTCAGAGCGGGGTCTCTCATC 3825
QY 1021 TTTTGAAGAGAGAGCGCGGAGCTTCAAAACACTGCTATTTGGCAATGTCTATGATTACAGAC 1080
```

Db 3826 TTTGGAGCAAGGCTCAGAGAAAACAAATGTGGACATTTGAAAGGTCTATGATTACAGAC 3885  
Qy 1081 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAGAGATTTGGGACCGTGGCAGTC 1140  
Db 3886 GAAGAGAAATCAGAACAACTCCGHTGGCTACGGACGATATGGTTCTGTATCTACC 3945  
Qy 1141 AATTTCCAGAGCAGCAGACAGACCTCTGCGACCGGAGATGTGCATGTGATGGAGCATTA 1200  
Db 3946 AACCTCCAGAGAGGCAACAGCAAGCAGCTACCGCAGATGTCAACACACAAGCGCTTCT 4005  
Qy 1201 CTGGCATGGTGGCAAGATAGAGACGTGTACTCGACGGTCCCATTTTGGGCCAAAATT 1260  
Db 4006 CCAGGCATGGTCTGGCAGACAGAGATGTGTACTTTACGGGGCCCATCTGGGCAAGATT 4065  
Qy 1261 CTTACACAGATGGACATTTTCAACCGTCTCTCTTTATGGGGCGCTTTGGACTCAAGAAC 1320  
Db 4066 CCACACCGAGCGGACATTTTCAACCCCTCTCCCTCATGGTGGATTCGGACTTAACAC 4125  
Qy 1321 CCGCCTCTCAGATCCTCATCAAAAACACGCCCTGTTCCTGCGAATCCTCCGGCGGATTT 1380  
Db 4126 CTTCTCTCCACAGATTTCTCATCAAGAACACCCCGGTACTTCCGAATCCTTCGACCACCTTC 4185  
Qy 1381 TCAGTACAAAATTTGCTTCATTCATCACCAATACTCCACAGGACAAAGTGTGGAA 1440  
Db 4186 AGTGGCGGAAATTTGCTTCTTTCATCACACAGTACTCCACGGGACAGGTTCAGCGTGGAG 4245  
Qy 1441 ATTGAATGGAGCTGCAGAAAGAAAACAGACGCGTGGAAATCCCGAAGTGCAGTACACA 1500  
Db 4246 ATCGATGGGAGCTGCAGAAAGGAAAACAGAACGCTGGAAATCCCGAAATTCAGTACACT 4305  
Qy 1501 TCCAAATATGCAAAATCTGCAACGTTGATTTTACTGTGGACAACAATGAGACTTTTATCT 1560  
Db 4306 TCCAACTACACAAGTCTGTATATGTGGACTTTACTGTGGACACTAATGGCGGTATTCA 4365  
Qy 1561 GAGCCTCGCCCATNTGGACCCGTTACTTACCCGTCCTCTGTA 1604  
Db 4366 GAGCCTCGCCCATNTGGACCCAGATACCTGACTCGTAATCTGTA 4409

## RESULT 9

US-09-770-315-3

; Sequence 3, Application US/09770315

; Patent No. 6429001

; GENERAL INFORMATION:

; APPLICANT: Chiron Corporation

; TITLE OF INVENTION: Recombinant AAV Packaging Systems

; FILE REFERENCE: 20263-501

; CURRENT APPLICATION NUMBER: US/09/770,315

; CURRENT FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: US 60/178,536

; PRIOR FILING DATE: 2000-01-26

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 7557

; TYPE: DNA

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: recombinant DNA

US-09-770-315-3

Query Match 64.0%; Score 1026.4; DB 3; Length 7557;  
Best Local Similarity 77.9%; Pred. No. 0;  
Matches 1250; Conservative 0; Mismatches 351; Indels 3; Gaps 1;

Qy 1 ATGGCTTCAGCGGTGGCGCACCACCAATGGCAGACAATAACGAAGCGCGCCGAGGTGGGT 60  
Db 2837 ATGGCTACAGGCTGGCGCACCAATGGCAGACAATAACGAAGCGCGCGGAGTGGGT 2896  
Qy 61 AATGCTCTCAGGAATTTGGCATTTGGCATTTCCATATGGCTTCACATGGCTGGCGACAGATCATCAC 120  
Db 2897 AATTCCTCGGGAATTTGGCATTTGGCATTTCCATATGGATGGCGGACAGATCATCACCA 2956

Db 4034 CCAGGCATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCCATCTCTGGCAAAAGATT 4093  
Qy 1261 CCTCACAGATGGACATTTTACCGGTCTCTTATGGGGGGCTTGGGACTCAAGAAC 1320  
Db 4094 CCACACACGAGCGGACATTTTACCCCTCTCCCTCATGGGTGGATTCGGACTTAAACAC 4153  
Qy 1321 CGCGCTCTCAGATCTCATCAAAACACGCTGTTCCTGCGAAATCCCTCCGGCGGAGTTT 1380  
Db 4154 CTCTCTCCACAGATTTCTCATGAAGAACCCCGGTACCTTGGAAATCTTGGACCACTTC 4213  
Qy 1381 TCAGCTACAAAGTTTGTCTTATTCATCACCCAATCTCCACAGGACAAAGTGTGGAA 1440  
Db 4214 AGTGGCGCAAGTTTGTCTTCTTCATCACAGACTCTCCACGGACAGGTTCAGCGTGGAG 4273  
Qy 1441 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500  
Db 4274 ATCGAGTGGAGCTGCAGAAAGGAAAAACAGCAACGCTGGAATCCCGAAATTCAGTACAT 4333  
Qy 1501 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTTATCT 1560  
Db 4334 TCCAACTACAAAGTCTGTATATGTGGACTTTTACTGTGGACACTAATGGCGTGTATTCAC 4393  
Qy 1561 GAGCCTCGCCCATTTGGCACCCGTTTACCTTACCCGTCCTCTGTA 1604  
Db 4394 GAGCCTCGCCCATTTGGCACCCAGATACCTGACTCGTAACTCTGTA 4437

RESULT 10  
US-09-770-315-2  
; Sequence 2, Application US/09770315  
; Patent No. 6429001  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Recombinant AAV Packaging Systems  
; FILE REFERENCE: 20263-501  
; CURRENT APPLICATION NUMBER: US/09/770,315  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,536  
; PRIOR FILING DATE: 2000-01-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 8698  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: recombinant DNA  
US-09-770-315-2

Query Match 64.0%; Score 1026.4; DB 3; Length 8698;  
Best Local Similarity 77.9%; Pred. No. 0;  
Matches 1250; Conservative 0; Mismatches 351; Indels 3; Gaps 1;

Qy 1 ATGGCTTACAGGCGGTGGCGCACCAATGGCAGACAAATAACGAAGGCGCCGACGAGTGGGT 60  
Db 2809 ATGGCTACAGGCGAGTGGCGCACCAATGGCAGACAAATAACGAAGGCGCCGACGAGTGGGT 2868  
Qy 61 AATGCTTCAGGAAATTTGGCATTCGGATTCCCATGCGCTGGCGACAGAGTCAATCACACC 120  
Db 2869 AATTCCTCGGAAATTTGGCATTCGGATTCCCATGCGCTGGCGACAGAGTCAATCACACC 2928  
Qy 121 AGCACCGGACCTGGGCGCTGGCCACTCAATTAACCACTCTACAGCAAAATCTTCAGT 180  
Db 2929 AGCACCGGAACTGGGCGCTGGCCACTCAATTAACCACTCTACAGCAAAATTTCCAGC 2988  
Qy 181 GCTTCAACGGGGCCAGCAACGACCACTACTCTGGGTACAGCACCCCTCTGGGGGTAT 240  
Db 2989 CAATCA---GGAGCTTCGAACGACATCACTATTTGGCTACAGCACCCCTTGGGGGTAT 3045  
Qy 241 TTTGATTCAACAGATTCCATGCCACTTTTCAACGAGTGGCAGCGACTCATCAAC 300  
Db 3046 TTTGACTTCAACAGATTTCACACTGCCACTTTTCAACGAGTGGCAGGACTCATCAAC 3105

Qy 301 AACAAATTGGGATTTCCGGSCCAAGAGACTCAACTTTCAAACCTCTTCAACATCAAGTCAAG 360  
Db 3106 AACAACTGGGATTTCCGAGCCCAAGAGACTCAACTTTCAAGCTCTTTAAATTTCAAGTCAAA 3165  
Qy 361 GAGGTACAGACGAATGTATGGCGTCACAAACATTCGCTATAAACCCTTTACAGACACGGTTCAA 420  
Db 3166 GAGGTACAGCAGAAATGACGGTACGAGCAGGATTCGCAATAAACCCTTTACAGACACGGTTCA 3225  
Qy 421 GTCTTCTCGGACTCGGAGTACAGCTTCGATACGTCCTCGGCTCTGCGGACCGAGGCTGC 480  
Db 3226 GTGTTTACTGACTCGGAGTACAGCTTCGATACGTCCTCGGCTCGGCGCATCAAGGATGC 3285  
Qy 481 CTCCCTCCGTTCCCGGCGGAGGTTCATGATTCGCGAATACGGCTACTCGACGCTCAAC 540  
Db 3286 CTCCCGCGGTTCCCGACGAGAGCTCTTCATGTGCGCACAGTATGGATACCTCACCCCTGAAC 3345  
Qy 541 AATGCAAGCCCAAGCGCTGGGACGTTTCATCTCTTTTACTCGCTCGGAATATTTCCCTTCTCAG 600  
Db 3346 AACGGAGTCAAGGAGTAGGACGCTCTTCATTTTACTCGCTGGAGTACTTTCTTCTCAG 3405  
Qy 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAAGTGCCTTTCCAC 660  
Db 3406 ATGCTGCGTACCGGAAAAACAATTTACCTTCAGCTACACTTTTGGAGGACGTTCTCTTCCAC 3465  
Qy 661 AGCAGCTACGGCGCACAGCGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720  
Db 3466 AGCAGCTACGCTCACAGCCAGAGTCTGGACCGTCTCATGAATCTCTCATCGACCGTAC 3525  
Qy 721 CTGTATTACTGAAACAGAACTCAAAATCAGTCCGGAAGTGGCCCAAAACAAGGACTTGTCTG 780  
Db 3526 CTGTATTACTTGAGCAGAAACAACAACACTCCAAAGTGGAAACCAACACGACGCTCAAGGCTTCAG 3585  
Qy 781 TTTAGCCGTGGTCTCCAGCTGGGATGTCTGTTTACGCCCAAAACAACCTGGCTACTTGGACCC 840  
Db 3586 TTTTCTCAGGCGCGGAGCGAGTGACATTCGGGACCAAGTCTAGGAACTGGCTCTCGGACCC 3645  
Qy 841 TGTTATCGGACGACGCGGTTTCTTAAACAACAAACAGACAAACAACAAGCAATTTTACC 900  
Db 3646 TGTTACCCCGCAGCAGCGAGTATCAAGAACAATCGCGGATTAACAACAGTGATACTGG 3705  
Qy 901 TGGACTGTGTCTTCAAAATATAAACCCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 960  
Db 3706 TGGACTGGAGCTTACCAGTACCACTCAATGGCAGAGACTCTCTGTGTAATCCGGGCGCG 3765  
Qy 961 GCTATGGCTCTACACAAAGACGACGAAAGATTTCTTTCCCATGAGCGGTGTCTATGATT 1020  
Db 3766 GCCATGGCAAGCCACAGGACGATGAAGAAAGTTTTTTTCTCTCAGAGCGGGTTCTCATC 3825  
Qy 1021 TTTTGGAAAGAGAGCGCGGAGCTTCAAAACACTGCAATTTGGACAATGTCTATGATTACAGAC 1080  
Db 3826 TTTTGGAAAGCAGGCTTCAGAGAAAACAATGTGGACATTTGAAAGGTATGATTATACAGAC 3885  
Qy 1081 GAAGAGGAAATTTAAAGCCACTAACCCCTGTGGCACCCGAAAGATTTGGGACCGTGGCAGTC 1140  
Db 3886 GAAGAGGAAATCAGGACAAACCAATCCCGTGGCTACGAGAGCATATGGTTCTGTATCTACC 3945  
Qy 1141 AATTTCCAGAGCAGCAGACACGCTTCGGACCGGAGATGTGCATGCTATATGGGAGCATTA 1200  
Db 3946 AACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAACACACAAGCGGTTCTT 4005  
Qy 1201 CCTGGCATGTGTGGCAAGATAGACGTTGTAACCTGACAGGTCCTCAATTTGGGCCCAAAAT 1260  
Db 4006 CAGGAGTGTGTGGCAGACAGAGATGTGTACCTTTCAGGGGCCCATCTTGGGCAAAAGATT 4065  
Qy 1261 CCTCACAGATAGGACACTTTTCCACCGCTCTCTCTTATTTGGGCGGCTTTGGACTCAAGAAC 1320  
Db 4066 CCACACAGGACGAGCAATTTTCCACCCCTCTCCCTCATGGGTGGATTCGAGCTTAAACAC 4125  
Qy 1321 CGGCTCTCAGATCTCTATCAAAAACAGCGCTGTTTCTTGGCAATCTCTCCGGCGGAGTTT 1380  
Db 4126 CCTCTCTCAACAGATTTCTATCAAGAAACACCCCGGTAACCTGCGAATCTTCTCGAACCCCTTC 4185









QY 1201 CCTGGCATGTTGGCAAGATAGACGTTGACCTGACGGTCCCATTTGGGCCAAATTT 1260  
DB |||||  
QY 4006 CAGGATGTTCTGGCAGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 4065  
DB |||||  
QY 1261 CCTCACAGATGGACATTTTCACCCGTCTCTCTTATATGGGGGGCTTTTGGACTCAAGAAC 1320  
DB |||||  
QY 4066 CCACACACGACGGACATTTTCACCCCTCTCCCTCATCGGTGGATTGGACTTAAACAC 4125  
DB |||||  
QY 1321 CGGCTCTCAGATCTCTCAATCAAAAACACGCTGTCTTCTGCGAATCTCCCGGGAGTTT 1380  
DB |||||  
QY 4126 CCTCTCCAAGATTTCTCATCAAGAACACCCCGGTACCTGCGAATCTCTTCGACCACCTTC 4185  
DB |||||  
QY 1381 TCAGCTACAAAGTTTCTTCAATTCATCCCAATCTCCACAGGACA-AGTCAGGTGGA 1439  
DB |||||  
QY 4186 AGTGGGCAAGTTTGTCTTCTTCATCAACAGTACTCCACGGGACACGGTCAGCGTGA 4245  
DB |||||  
QY 1440 AATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACAC 1499  
DB |||||  
QY 4246 GATCGAGTGGAGCTGCAGAGGAAAACAGCAACGCTGGAATCCGGAATTCAGTACAC 4305  
DB |||||  
QY 1500 ATCCAAATTAAGCAAAATCTGCCAAGCTTGAATTTTACTGTGGACAACAATGGAGTTTATAC 1559  
DB |||||  
QY 4306 TTCCAATCAACAAGCTGTGTTAATCGTGGACTTACCGTGGATCTAATGGCGTGTATTC 4365  
DB |||||  
QY 1560 TGAGCTCGCCCATTTGGCACCCTGTACCTTACCGTCCCTGTA 1604  
DB |||||  
QY 4366 AGAGCCTCGCCCATTTGGCACCAGATACCTGACTCGTAATCTGTA 4410  
DB |||||

RESULT 14

US-08-709-609-1  
/ Sequence 1, Application US/08709609  
/ Patent No. 5858775  
/ GENERAL INFORMATION:  
/ APPLICANT: Johnson, Philip R.  
/ TITLE OF INVENTION: Adeno-Associated Virus Materials and  
/ TITLE OF INVENTION: Methods  
/ NUMBER OF SEQUENCES: 3  
/ CORRESPONDENCE ADDRESS:  
/ ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
/ STREET: 6300 Sears Tower, 233 S. Wacker Drive  
/ CITY: Chicago  
/ STATE: Illinois  
/ COUNTRY: USA  
/ ZIP: 60606  
/ COMPUTER READABLE FORM:  
/ MEDIUM TYPE: Floppy disk  
/ COMPUTER: IBM PC compatible  
/ OPERATING SYSTEM: PC-DOS/MS-DOS  
/ SOFTWARE: PatentIn Release #1.0, Version #1.25  
/ CURRENT APPLICATION DATA:  
/ FILING DATE: US/08/709,609  
/ CLASSIFICATION: 435  
/ ATTORNEY/AGENT INFORMATION:  
/ NAME: No. 5858775and, Greta E.  
/ REGISTRATION NUMBER: 35,302  
/ REFERENCE/DOCKET NUMBER: 31975  
/ TELECOMMUNICATION INFORMATION:  
/ TELEPHONE: (312) 474-6300  
/ TELEFAX: (312) 474-0448  
/ TELEX: 25-3856  
/ INFORMATION FOR SEQ ID NO: 1:  
/ SEQUENCE CHARACTERISTICS:  
/ LENGTH: 4680 base pairs  
/ TYPE: nucleic acid  
/ STRANDEDNESS: single  
/ TOPOLOGY: linear  
/ MOLECULE TYPE: DNA (genomic)  
US-08-709-609-1  
Query Match 62.6%; Score 1004.2; DB 2; Length 4680;

Best Local Similarity 77.4%; Pred. No. 0;  
Matches 1243; Conservative 0; Mismatches 358; Indels 4; Gaps 2;  
QY 1 ATGGCTTCAGCGGTGGCGCACTTAATGGCAGCAATAAAGAAAGCGCCGACGGAGTGGGT 60  
DB |||||  
QY 2809 ATGGCTACAGCGAGTGGCGCACTTAATGGCAGCAATAAAGAAAGCGCCGACGGAGTGGGT 2868  
DB |||||  
QY 61 AATGCTTCAGAAATTTGGCATTTGGCATTTCCACATGGCTGGGGCGACAGAGTATCAACACC 120  
DB |||||  
QY 2869 AATTCCTCCGAAATTTGGCATTTGGCATTTCCATATGGATGGGGCGACAGAGTATCAACACC 2928  
DB |||||  
QY 121 AGCACCCGCACTTGGGCTTGGCCACCTTCAATTAACCACTCTCTCAAGCAAAATCTCCAGT 180  
DB |||||  
QY 2929 AGCACCCGCACTTGGGCTTGGCCACCTTCAATTAACCACTCTCTCAAGCAAAATTTCCAGC 2988  
DB |||||  
QY 181 GCTTCAACGGGGGCGCAGCAACGACATCTTTCGGCTACAGACCCCTGGGGGTAT 240  
DB |||||  
QY 2989 CAATCA---GGAGCCTCGAAGCAATCACTACTTTGGCTACAGCACCCCTTGGGGGTAT 3045  
DB |||||  
QY 241 TTTGATTTCAACAGATTTCCATCTGCCACTTTTCAACAGTGTGGCGAGTCTCATCAAC 300  
DB |||||  
QY 3046 TTTGATTTCAACAGATTTCCATCTGCCACTTTTCAACAGTGTGGCGAGTCTCATCAAC 3105  
DB |||||  
QY 301 AACAAATGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCAAGTCAAG 360  
DB |||||  
QY 3106 AACAACTGGGATTTCCGACCCCAAGAGACTCAACTTTCAAGCTTTTAACTTCAAGTCAAA 3165  
DB |||||  
QY 361 GAGGTCAAGCAATGATGGCGTCAACACATCTCGCTAATAAATCTTACAGACAGTGTCAA 420  
DB |||||  
QY 3166 GAGGTCAAGCAATGATGGCGTCAACACATCTCGCTAATAAATCTTACAGACAGTGTCAA 3225  
DB |||||  
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCCGTACGCTCTCGGCTCTGCGACACAGGGTGC 480  
DB |||||  
QY 3226 GTGTTTACTGACTCGGAGTACAGCTTCCGTACGCTCTCGGCTCGGCGCATCAAGATGC 3285  
DB |||||  
QY 481 CTCCCTCTCGTTCCTCCGGCGAGCGTGTTCATGATTTCCGCAATACGGTACTCTGACGCTCAAC 540  
DB |||||  
QY 3286 CTCCCGCGGTTCCTCCAGCAGACAGCTTTCATGGTGGCCACAGTATGGATACCTCACCCCTGAAC 3345  
DB |||||  
QY 541 AATGCGACGCAAGCGTGGAGTGTATCTCTTTTACTGCTGGAAATATTTTCCCTCTCAG 600  
DB |||||  
QY 3346 AACGGGAGTACGAGCAGTAGGACGCTCTTCAATTTACTGCTGGAGTACTTCTCTCTCAG 3405  
DB |||||  
QY 601 ATGCTGAGAAAGGGCAACAACTTTTACCTTACGCTACACCTTTGAGGAAGTGTCTTCCAC 660  
DB |||||  
QY 3406 ATGCTGCGTACCGGAAACAACTTTTACCTTACGCTACACCTTTTGGAGAGTGTCTTCCAC 3465  
DB |||||  
QY 661 AGCAGCTACGCGCACAGCAGCGCTGGAACGGGTGATGAATCTCTCTCATCGACCAATAC 720  
DB |||||  
QY 3466 AGCAGCTACGCTACAGCAGCAGAGTCTGGACCGTCTCATGAATCTCTCATCGACCAATAC 3525  
DB |||||  
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGGCCCAAAACAGAGACTTGTG 780  
DB |||||  
QY 3526 CTGTATTACTTGAGCAGAAACAACTCAAGTGGAAACCAACCCAGCAGTCAAGGCTTCAG 3585  
DB |||||  
QY 781 TTTAGCCGTGGGTCTCGAGCTGGCATGTCTGTTGAGCCCAAAACCTGCTACTCGACCC 840  
DB |||||  
QY 3586 TTTTCTCAGGCGGAGCGAGTGACATTCGGGACCAAGTCTAGGAATCTGGCTTCTCGGACCC 3645  
DB |||||  
QY 841 TGTATTACGCGCAGCGGTGTTTCTTAAACAAAAACAGACAAACAAACAGCAATTTTACC 900  
DB |||||  
QY 3646 TGTATTACGCGCAGCGGAGTATCAAGGATCTGCGGATTAACAAACAGTGAATACG 3705  
DB |||||  
QY 901 TGGACTGTGCTTCAAAATATATACCTCAATGGGCTGGAATCCATCATCAACCCCTGGCACT 960  
DB |||||  
QY 3706 TGGACTGGAGCTTACCAAGTACCACTCAATGGCAGAGACTCTCTGGTGAATCCGGGGCCC 3765  
DB |||||  
QY 961 GCTATGGCTCTCACAAAGACGACAGAGTGTCTTTCCCATGAGCGGTGTCTATGATT 1020  
DB |||||  
QY 3766 GCCATGGCAAGCCCAAGGACGATGAAGAAAAAGTTTTTCTCAGAGCGGGTTCATC 3825  
DB |||||  
QY 1021 TTTGAAAGAGAGCGCGGAGCTTCAAACTGCAATTTGGCAAAATGTCTATGATTACAGAC 1080  
DB |||||



3826	TTTTGGGAGCAAGCTCAGAGAAAAA	CAATGTGGA	CAATTGAAAAGGTCATGATTACAGAC	3885
1081	GAAGAGGAAATTAAGCCACTAAAC	CTCTGTGGCCACCGAAAGATTTGGGA	CCGTGGCAGTC	1140
3886	GAAGAGGAATCGGAACA	CCAATCCCGTGGCTACGGAGCAGTATGGTCTGTATCTTACC	3945	
1141	AATTTCCAGAGCAGCAGCA	CAGACCTTCGCGACCGGAGATGTGCATGTGGGAGCAATTA	1200	
3946	AACTTCAGAGAGGCAACAGACA	AGCAGCTACCGCAGATGTCAACA	CACAAGCGCTTCTT	4005
1201	CTTGGCATGTGTGGCAAGATAGAG	CGTCTACTCGCAGGTCCTCATTTGGGCCCAAAATT	1260	
4006	CCAGGCATGCTGGCGAGGACAGAT	GTGTACTCTCAGGGGGCCATCTGGGCAAAAGATT	4065	
1261	CTTCACACAGATGGACATTTT	CACCCGCTCTCTTATGGCGGCTTTGACCTCAAGAAC	1320	
4066	CCACACACGACGGA	CATTTTCAACCCCTCTCCCTCATGGGTGGATTCGACCTTAAACAC	4125	
1321	CCGCTCTCTCAGATCTCATCA	AAAAACA	CGCTGTTCCTCGGAATCTCCGGCGGAGTTT	1380
4126	CTCTCTCCA	CAGATCTCATCAAGAA	CACCCCGGTACTCTCGGAATCTTCGACCACTTC	4185
1381	TCAGCTACAAAGTTTGCTTCAT	TTCATCA	CCCAATATCCACAGACA-AGTAGTGTGGA	1439
4186	AGTCGGCAAGTTTGCTTCTT	CATCA	CAGTACTCCACGGGACACGGCTGACGCTGGA	4245
1440	AATTTGAATGGAGCTGCAGAA	AGAAACACGACGCTGGAA	TCCCGAAGTCAGATACAC	1499
4246	GATCGAGTGGGAGCTCAGAA	AGGAAACACGCA	AAACGCTGGAAATCCCGAAATTCAGTACAC	4305
1500	ATCCAAATTATGC	AAAAATCTGCAACGTTGATTTT	TACTGTGGACAACAATGGACTTTATAC	1559
4306	TTCCAACTACA	CAAGTCTGTTAATCTGTGAC	TACTACCTGGATTAATGGCGTGTATTC	4365
1560	TGAGCCTCGCCCCAT	TGGCCACCGTTTACCTTAC	CCGTCCTCCCTGTA	1604
4366	AGAGCTCGCCCCAT	TGGCACCAAGATACCTTGACTCGTA	AACTCTGTA	4410

RESULT 15

<p>PCT-US95-07178-1</p> <p>; Sequence 1, Application PC/TUS9507178</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Johnson, Philip R.</p> <p>; TITLE OF INVENTION: Adeno-Associated Virus Materials and</p> <p>; TITLE OF INVENTION: Methods</p> <p>; NUMBER OF SEQUENCES: 3</p> <p>; CORRESPONDENCE ADDRESS:</p> <p>; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &amp; Borun</p> <p>; STREET: 6300 Sears Tower, 233 S. Wacker Drive</p> <p>; CITY: Chicago</p> <p>; STATE: Illinois</p> <p>; COUNTRY: USA</p> <p>; ZIP: 60606</p> <p>; COMPUTER READABLE FORM:</p> <p>; MEDIUM TYPE: Floppy disk</p> <p>; COMPUTER: IBM PC compatible</p> <p>; OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>; SOFTWARE: Patent in Release #1.0, Version #1.25</p> <p>; CURRENT APPLICATION DATA:</p> <p>; APPLICATION NUMBER: PCT/US95/07178</p> <p>; FILING DATE:</p> <p>; CLASSIFICATION:</p> <p>; ATTORNEY/AGENT INFORMATION:</p> <p>; NAME: Noland, Greta E.</p> <p>; REGISTRATION NUMBER: 35,302</p> <p>; REFERENCE/DOCKET NUMBER: 31975</p> <p>; TELECOMMUNICATION INFORMATION:</p> <p>; TELEPHONE: (312) 474-6300</p> <p>; TELEFAX: (312) 474-0448</p> <p>; TELEX: 25-3856</p> <p>; INFORMATION FOR SEQ ID NO: 1:</p> <p>; SEQUENCE CHARACTERISTICS:</p>	<p>3226 GTGTTTACTGACTCGAGTACACGCTCCGTAAGTCTCTCGCTCGGCGCATCAAGGATGC 3285</p> <p>481 CTCCTCCGTTCCCGGCGAGTGTTCATGATTCGGAAATACGCTACCTGACGCTCAAC 540</p> <p>3286 CTCCGCGGTTCCACAGCAGAGCTCTTCATGTGTGCGACAGTATGATACCTCACCTCAAC 3345</p> <p>541 AATGCGACCGAAGCCGTGCGAGCTTCACTCTTTTACTGCCTTGGAAATATTTCCCTTCTTCAG 600</p> <p>3346 AACGGAGTCAAGCAGTAGGACGCTCTTCATTTTACTGCTTGGAGTACTTTCCTTCTTCAG 3405</p> <p>601 ATGCTGAGAACGGGCAACAACCTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660</p> <p>3406 ATGCTGCGTACCGGAAACAACCTTTACCTTCAGCTACACCTTTTGAAGGACGTTCTCTTTCCAC 3465</p> <p>661 AGCAGTACCGGCACAGCCAGAGCTGACCGGTGATGAATCTCTCATCGACCAATAC 720</p> <p>3466 AGCAGCTACGCTCACAGCCAGAGTCTGACCGCTCTCATGAATCTCTCATCGACCAATAC 3525</p> <p>721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAGGACTTCTGT 780</p> <p>3526 CTGTATTACTTGAGCAGAAACAACACTCCAAGTGGAAACCAACCGCAGTCAAGGCTTCAG 3585</p> <p>781 TTTAGCGGTGGTCTCCAGCTGGCATGTCTGTTGAGCCCAAAAATCGCTACTCTGGACCC 840</p> <p>3586 TTTTCTCAGGCCGAGCGAGTGACATTCGCGGACCAAGTCTAGGAACTGGCTTCTGGGACCC 3645</p> <p>841 TGTATTGCGCAGCAGCGGTTTCTAAAAACAAAAACAGACAAACAACAGCAATTTTACC 900</p> <p>3646 TGTATCCGCCAGCAGCGAGTATCAAGAAGATCTCGGATACACACACAGTGAATACTCG 3705</p> <p>901 TGGACTGGTGCTTCAAAAATAATACCTCAATGGGCGTGAATCCATCATCAACCCCTGGCACT 960</p>
--	--

```
Db 3706 TGGAGCTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGGGCCC 3765
Qy 961 GCTATGGCCTCACACAAAGACGACGAAGATTCTTTCCCATGAGCGGTGTCATGATT 1020
Db 3766 GCCATGGCAAGCCACAAGGACGATGAAGAAAGTTTTTTCTCTCAGAGCGGGTTCTCATC 3825
Qy 1021 TTTGGAAAGAGAGCGCGGAGCTTCAACACCTGCTATGGCAATGTCTATGATTACAGAC 1080
Db 3826 TTTGGAAAGCAAGGCTCAGAGAAAACAATGTGAACATTTGAAAAGGTCAATGATTACAGAC 3885
Qy 1081 GAAGAGGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140
Db 3886 GAAGAGGAAATCGGAACCAACCAATCCCGTGGCTACGGAGCAGTATGGTTCTGTATCTACC 3945
Qy 1141 AATTTCAGAGCAGCAGCAGACCCCTGGGACCGGAGATGTGCATGCTATGGGAGCATTA 1200
Db 3946 AACCTCCAGAGAGGGAACAAGCAGCTACCGCAGATGTCAACACACAAGCGCTTCTT 4005
Qy 1201 CCTGGCATGGTGTGCAAGATAGAGACGTGTACTGACGGGTCCCATTTGGGCCAAAAATT 1260
Db 4006 CCAGCATGGTCTGGCAGACAGAGATGTGTACCTTCAGGGGCCCATCTGGGCAAGATT 4065
Qy 1261 CCTCACAGATGGACACTTTTCAACCGCTCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 1320
Db 4066 CCACACGGACGGACATTTTCAACCCCTCTCCCTCATGGGTGGATTGGACTTAAACAC 4125
Qy 1321 CCGCCTCCTCAGATCCTCATCAAAAACAGCGCTGTCTTCGGAATCTCCGGCGGAGTTT 1380
Db 4126 CCTCCTCCACAGATTCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACCACCTTC 4185
Qy 1381 TCAGCTACAAAGTTTGTCTTTCATTTCATCCCAATCTCCACAGGACA-AGTGAGTGTGA 1439
Db 4186 AGTGGGCAAGTTTGTCTTCTTCATCACAGTACTCCACGGGACCGTCAGCGTGA 4245
Qy 1440 AATTGAATGGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCGGAAGTGCAGTACAC 1499
Db 4246 GATCGAGTGGAGCTGCAGAGGAAACAGCAACGCTGGAATCCGGAATTCAGTACAC 4305
Qy 1500 ATCCAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGGACTTTATAC 1559
Db 4306 TTCCAACACTACAACAAGTCTGTAAATCGTGGACTTACCGTGGATACTAATGGCGGTGTTTC 4365
Qy 1560 TGAGCCTCGCCCATTTGGCACCCGTTTACCCTTACCCGTCCTCCCTGTA 1604
Db 4366 AGAGCCTCGCCCATTTGGCACCAAGATACCTGACTGTGAATCTGTA 4410
```

Search completed: November 29, 2005, 14:27:50  
Job time : 233.137 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: November 28, 2005, 08:46:14 ; Search time 750.542 Seconds  
(without alignments)  
14243.261 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_2829\_4432  
Perfect score: 1604  
Sequence: 1 atgggttcaggcggggggc.....taccttaccgtccctcgtga 1604

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues  
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_21.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1604	100.0	1605	3 AAD00779	Ad00779 Adeno-ass
2	1604	100.0	1800	3 AAD00778	Ad00778 Adeno-ass
3	1604	100.0	2211	3 AAD00777	Ad00777 Adeno-ass
4	1604	100.0	2211	14 AD227052	Ad227052 Adeno-ass
5	1604	100.0	2211	14 AD226929	Ad226929 Adeno-ass
6	1604	100.0	4347	13 ADW39398	Adw39398 Adeno-ass
7	1604	100.0	4347	14 AD246594	Ad246594 HSV-AAV s
8	1604	100.0	4718	3 AAD00772	Ad00772 Adeno-ass
9	1604	100.0	4718	10 AD276507	Ad276507 Adeno-ass
10	1604	100.0	4718	10 ADL13984	Adl13984 Adeno-ass
11	1604	100.0	4718	12 ADG39758	Adg39758 AAV-1 gen
12	1604	100.0	7447	8 ACC58477	Acc58477 Vector pf
13	1600.8	99.8	2211	14 AD226931	Ad226931 Adeno-ass
14	1594.4	99.4	2214	14 AD226930	Ad226930 Adeno-ass
15	1592.8	99.3	2211	14 AD226928	Ad226928 Adeno-ass
16	1591.2	99.2	2211	14 AD226932	Ad226932 Adeno-ass
17	1512.8	94.3	2211	14 AD227053	Ad227053 Adeno-ass
18	1512.8	94.3	4239	13 ADW39402	Adw39402 Adeno-ass
19	1512.8	94.3	4239	14 AD246598	Ad246598 HSV-AAV s

20	1512.8	94.3	4683	4 AAF23749	Aaf23749 AAV6 DNA
21	1512.8	94.3	4683	10 ADL13983	Adl13983 Adeno-ass
22	1512.8	94.3	4683	12 ADG39763	Adg39763 AAV-6 gen
23	1104.6	68.9	2214	14 ADZ26890	Adz26890 Adeno-ass
24	1084	67.6	2214	14 ADZ26891	Adz26891 Adeno-ass
25	1079	67.3	2214	14 ADZ26886	Adz26886 Adeno-ass
26	1077.4	67.2	2214	14 ADZ26885	Adz26885 Adeno-ass
27	1075.8	67.1	2217	14 ADZ26892	Adz26892 Adeno-ass
28	1074.2	67.0	2214	14 ADZ26884	Adz26884 Adeno-ass
29	1074.2	67.0	2214	14 ADZ26882	Adz26882 Adeno-ass
30	1072.6	66.9	2214	14 ADZ26889	Adz26889 Adeno-ass
31	1071.6	66.8	3094	10 ADE76523	Ad276523 Adeno-ass
32	1071	66.8	2214	14 ADZ26883	Adz26883 Adeno-ass
33	1070	66.7	2214	14 ADZ26887	Adz26887 Adeno-ass
34	1070	66.7	2214	14 ADZ27063	Adz27063 Adeno-ass
35	1070	66.7	3095	10 ADE76525	Ad276525 Adeno-ass
36	1070	66.7	3095	10 ADE76524	Ad276524 Adeno-ass
37	1070	66.7	4721	10 ADE76502	Ad276502 Adeno-ass
38	1070	66.7	4721	12 ADG39764	Adg39764 AAV-7 gen
39	1070	66.7	4721	14 ADV67509	Adv67509 Nucleotid
40	1068.4	66.6	4721	14 ADZ27030	Adz27030 Adeno-ass
41	1068.4	66.6	3127	10 ADE76516	Ad276516 Adeno-ass
42	1068	66.6	2209	14 ADZ26898	Adz26898 Adeno-ass
43	1068	66.6	3075	10 ADE76527	Ad276527 Adeno-ass
44	1066.4	66.5	2208	14 ADZ26895	Adz26895 Adeno-ass
45	1064.8	66.4	2208	14 ADZ26897	Adz26897 Adeno-ass

ALIGNMENTS

RESULT 1  
AAD00779  
ID AAD00779 standard; DNA; 1605 BP.  
XX AAD00779;  
XX  
XX 08-SEP-2000 (first entry)  
XX Adeno-associated virus serotype 1 capsid protein VP3 DNA.  
XX  
XX Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;  
KW cap protein; recombinant viral vector; gene delivery; gene therapy;  
KW vaccine; transgene; VP3; ss.  
XX  
XX Adeno associated virus serotype 1.  
XX  
XX Key Location/Qualifiers  
FT CDS 1..1605  
FT /\*tag= a  
FT /product= "VP3 protein"  
XX  
XX WO200028061-A2.  
XX  
XX 19-MAY-2000.  
XX  
XX 02-NOV-1999; 99WO-US025694.  
XX  
XX 05-NOV-1998; 98US-0107114P.  
XX  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX  
XX Wilson JW, Xiao W;  
XX WPI; 2000-376571/32.  
XX P-PSDB; AAY71169.  
XX  
XX Novel adeno-associated virus serotype 1 polynucleotide useful for  
XX preparation of medicament for delivery of a transgene to a host.  
XX Claim 10; Page 96-99; 108pp; English.  
XX  
XX The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA

which is characterised by two inverted terminal repeats (ITR) and open reading frames for rep and capsid (cap) proteins. The rep reading frame encodes four proteins, Rep 78, Rep 52 and Rep 40, while the cap reading frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1 sequence or its fragments particularly ITRs, rep and cap coding regions, are useful in production of recombinant viral vectors for gene delivery. These vectors can be used as gene therapy vectors, vaccine vectors or antisense delivery vectors. The AAV-1 does not induce the formation of neutralising antibodies specific to any serotype of AAV hence is useful for transforming host cells, and in the preparation of a medicament for the delivery of transgene to a host. The present sequence is an AAV-1 DNA encoding a cap protein VP3 which is useful in the production of recombinant viral vector for gene delivery

XX Sequence 1605 BP; 431 A; 466 C; 361 G; 347 T; 0 U; 0 Other;

Query Match		100.0%;	Score 1604;	DB 3;	Length 1605;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1604;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGCTTCAGGGCGTGGCGCAATATGGCAGACAATAACGAAGGCGCCGACGGAGTGGGT	60		
DB	1	ATGGCTTCAGGGCGTGGCGCAATATGGCAGACAATAACGAAGGCGCCGACGGAGTGGGT	60		
QY	61	AATGCTCAGGAAATGGCATTCGGAATCCAGATGCTGGGCGACAGAGTCATCACCAC	120		
DB	61	AATGCTCAGGAAATGGCATTCGGAATCCAGATGCTGGGCGACAGAGTCATCACCAC	120		
QY	121	AGCACCGCACCTGGGCGCTTGGCCACTACATAACCACTCTTACAAAGCAAAATCTCCAGT	180		
DB	121	AGCACCGCACCTGGGCGCTTGGCCACTACATAACCACTCTTACAAAGCAAAATCTCCAGT	180		
QY	181	GCTTCAACGGGGCCAGCAACGACAACTACTTCCGGCTACAGCAACCCCTGGGGGGTAT	240		
DB	181	GCTTCAACGGGGCCAGCAACGACAACTACTTCCGGCTACAGCAACCCCTGGGGGGTAT	240		
QY	241	TTTGATTTCAACAGATTCACATGCGCACTTTTACACAGTGAGCTGGCAGGAGCTCATCAAC	300		
DB	241	TTTGATTTCAACAGATTCACATGCGCACTTTTACACAGTGAGCTGGCAGGAGCTCATCAAC	300		
QY	301	AACAAATGGGGATTCGGGCGCCAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG	360		
DB	301	AACAAATGGGGATTCGGGCGCCAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG	360		
QY	361	GAGGTCAAGCAAGTAATGGGCTCAACATCGCTAAATACCTTACAGCAACGGTTCAA	420		
DB	361	GAGGTCAAGCAAGTAATGGGCTCAACATCGCTAAATACCTTACAGCAACGGTTCAA	420		
QY	421	GTCTTCTCGGACTCGGAGTACAGCTTCGGTAGCTCTCGGCTCTGGCAGCAGGCTGC	480		
DB	421	GTCTTCTCGGACTCGGAGTACAGCTTCGGTAGCTCTCGGCTCTGGCAGCAGGCTGC	480		
QY	481	CTCCCTCCGTTCCGCGCGAGCTGTTTCATGATTCGCAATACGGCTACCTGACGCTCAAC	540		
DB	481	CTCCCTCCGTTCCGCGCGAGCTGTTTCATGATTCGCAATACGGCTACCTGACGCTCAAC	540		
QY	541	AATGGCAGCAAGCGTGGGACGTTTCATCTTTTACTGCTTGAATATTTCCCTTCTCAG	600		
DB	541	AATGGCAGCAAGCGTGGGACGTTTCATCTTTTACTGCTTGAATATTTCCCTTCTCAG	600		
QY	601	ATGCTGAGACGGGCAACACTTTTACCTTCAGTACACCTTTGAGGAAGTGCCTTTCCAC	660		
DB	601	ATGCTGAGACGGGCAACACTTTTACCTTCAGTACACCTTTGAGGAAGTGCCTTTCCAC	660		
QY	661	AGCAGCTAGCGCACAGCAGAGCTGGACCGGCTGATGAATCTCTCATCGACCAATATAC	720		
DB	661	AGCAGCTAGCGCACAGCAGAGCTGGACCGGCTGATGAATCTCTCATCGACCAATATAC	720		
QY	721	CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAGGACTTGTCTG	780		
DB	721	CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAGGACTTGTCTG	780		
QY	781	TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTACCTGGACCC	840		

DB	781	TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTACCTGGACCC	840		
QY	841	TGTTATCGCAGCAGCGGTTTCTTAAACAAAAACAGACAAACAAACAGCAATTTTACC	900		
DB	841	TGTTATCGCAGCAGCGGTTTCTTAAACAAAAACAGACAAACAAACAGCAATTTTACC	900		
QY	901	TGGACTGGTCTTCAAAATATACTCAATGGGCGTGAATCCATCATCAACCCCTGGCACT	960		
DB	901	TGGACTGGTCTTCAAAATATACTCAATGGGCGTGAATCCATCATCAACCCCTGGCACT	960		
QY	961	GCTATGGCTCTACACAAAGACAGCAAGTCTTCTTCCATGAGCGGTGTCATGATT	1020		
DB	961	GCTATGGCTCTACACAAAGACAGCAAGTCTTCTTCCATGAGCGGTGTCATGATT	1020		
QY	1021	TTTGGAAAGAGAGCGCGGAGCTTCAAACTGCAATGGACAAATGTCTATGATTAACAGAC	1080		
DB	1021	TTTGGAAAGAGAGCGCGGAGCTTCAAACTGCAATGGACAAATGTCTATGATTAACAGAC	1080		
QY	1081	GAAGAGGAAATTAAGGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC	1140		
DB	1081	GAAGAGGAAATTAAGGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC	1140		
QY	1141	AATTTCCAGAGCAGCAGCACAGACCTCGGACCGGAGATGTGCATGTATGGGACATTA	1200		
DB	1141	AATTTCCAGAGCAGCAGCACAGACCTCGGACCGGAGATGTGCATGTATGGGACATTA	1200		
QY	1201	CCTGCACTGGTGTGGCAAGATAGAGACGTGTACTGCGAGGGTCCCATTTGGGCCAAAATT	1260		
DB	1201	CCTGCACTGGTGTGGCAAGATAGAGACGTGTACTGCGAGGGTCCCATTTGGGCCAAAATT	1260		
QY	1261	CCTCACAGATGGACACTTTTCAACCGCTCTCTTATGGGCGGCTTTGGACTCAAGAAC	1320		
DB	1261	CCTCACAGATGGACACTTTTCAACCGCTCTCTTATGGGCGGCTTTGGACTCAAGAAC	1320		
QY	1321	CGCGCTCTCGAGATCTCATCAAAAACACGCTGTCTCTGCGAATCTCTCGGCGGAGTTT	1380		
DB	1321	CGCGCTCTCGAGATCTCATCAAAAACACGCTGTCTCTGCGAATCTCTCGGCGGAGTTT	1380		
QY	1381	TCAGCTACAAAGTTTGTCTTCAATCCCAATATCTCCACAGGACAAAGTGTGGAA	1440		
DB	1381	TCAGCTACAAAGTTTGTCTTCAATCCCAATATCTCCACAGGACAAAGTGTGGAA	1440		
QY	1441	ATTGAATGGAGCTCGAGAAAGAAACAGACGCGCTGGAATCCCGAAGTGCAGTACACA	1500		
DB	1441	ATTGAATGGAGCTCGAGAAAGAAACAGACGCGCTGGAATCCCGAAGTGCAGTACACA	1500		
QY	1501	TCCAAATTTGCAAAATCTGCCAACTGATTTTACTGTGGACAAACATGGACTTTTATCT	1560		
DB	1501	TCCAAATTTGCAAAATCTGCCAACTGATTTTACTGTGGACAAACATGGACTTTTATCT	1560		
QY	1561	GAGCCTCGCCCATTTGGCACCCGTTTACCTTACCCCTGCCCTGTGA	1604		
DB	1561	GAGCCTCGCCCATTTGGCACCCGTTTACCTTACCCCTGCCCTGTGA	1604		

RESULT 2

ID	AAD00778
XX	AAD00778 standard; DNA; 1800 BP.
AC	AAD00778;
XX	
DT	08-SEP-2000 (first entry)
XX	
DE	Adeno-associated virus serotype 1 capsid protein VP2 DNA.
XX	
KW	Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;
KW	cap protein; recombinant viral vector; gene delivery; gene therapy;
XX	vaccine; transgene; VP2; ss.
OS	Adeno associated virus serotype 1.
XX	
FH	Key Location/Qualifiers

CDS 1..1800  
FT FT /\*tag= a  
FT FT /product= "VP2 protein"  
FT FT /partial  
XX  
PN WO200028061-A2.  
PD 18-MAY-2000.  
XX 02-NOV-1999; 99WO-US025694.  
XX 05-NOV-1998; 98US-0107114P.  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX PA  
XX Wilson JM, Xiao W;  
XX WPI; 2000-376571/32.  
DR P-PSDB; AAY71168.  
XX Novel adeno-associated virus serotype 1 polynucleotide useful for  
PT preparation of medicament for delivery of a transgene to a host.  
XX  
PS Claim 10; Page 90-93; 109pp; English.  
XX The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA  
CC which is characterised by two inverted terminal repeats (ITR) and open  
CC reading frames for rep and capsid (cap) proteins. The rep reading frame  
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap  
CC reading frame encodes three structural proteins, VP1, VP2 and VP3. The  
CC AAV-1 sequence or its fragments particularly ITRs, rep and cap coding  
CC regions, are useful in production of recombinant viral vectors for gene  
CC delivery. These vectors can be used as gene therapy vectors, vaccine  
CC vectors or antisense delivery vectors. The AAV-1 does not induce the  
CC formation of neutralising antibodies specific to any serotype of AAV  
CC hence is useful for transforming host cells, and in the preparation of a  
CC medicament for the delivery of transgene to a host. The present sequence  
CC is an AAV-1 DNA encoding a cap protein VP2 which is useful in the  
CC production of recombinant viral vector for gene delivery  
XX  
SQ Sequence 1800 BP; 483 A; 532 C; 409 G; 376 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1604; DB 3; Length 1800;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTTCAGCGGTGGCGCAACAAATGGCAGACAATAACGAAGCGCGGAGTGGGT 60  
DB 196 ATGGCTTCAGCGGTGGCGCAACAAATGGCAGACAATAACGAAGCGCGGAGTGGGT 255  
QY 61 AATGCTTCAGGAATTTGGCAATTCGATTCACATGCTGGCGCACAGAGTCATCACCACC 120  
DB 256 AATGCTTCAGGAATTTGGCAATTCGATTCACATGCTGGCGCACAGAGTCATCACCACC 315  
QY 121 AGCACCGGACCTGGGCTTGGCCCACTCAATAAACCACTCTACAGCAAAATCTCCAGT 180  
DB 316 AGCACCGGACCTGGGCTTGGCCCACTCAATAAACCACTCTACAGCAAAATCTCCAGT 375  
QY 181 GCTTCAACGGGGCCAGCAACAGCAACCACTACTCTGGGTACAGCACCCCTGGGGGTAT 240  
DB 376 GCTTCAACGGGGCCAGCAACAGCAACCACTACTCTGGGTACAGCACCCCTGGGGGTAT 435  
QY 241 TTTGATTTCAACAGATTCGATCGCACTTTTCCACGATGCTGCGGACGACCTCATCAAC 300  
DB 436 TTTGATTTCAACAGATTCGATCGCACTTTTCCACGATGCTGCGGACGACCTCATCAAC 495  
QY 301 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCCAAGTCAAG 360  
DB 496 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCCAAGTCAAG 555  
QY 361 GAGGTACGACGAATGATGGCGTCACAAACCATCGCTAATTAACCTTACGACGAGTTCAA 420  
DB 556 GAGGTACGACGAATGATGGCGTCACAAACCATCGCTAATTAACCTTACGACGAGTTCAA 615

QY 421 GTCTTCTCGGACTCGGAGTACCCAGCTTCGGTACGTCCTCGGCTCTGGCGACAGGGCTGC 480  
DB 616 GTCTTCTCGGACTCGGAGTACCCAGCTTCGGTACGTCCTCGGCTCTGGCGACAGGGCTGC 675  
QY 481 CTCCCTCCGTTCCCGCGGACGCTTCATGATTTCCGCAATACGGCTTACGCTCAAC 540  
DB 676 CTCCCTCCGTTCCCGCGGACGCTTCATGATTTCCGCAATACGGCTTACGCTCAAC 735  
QY 541 AATGCGAGCAAGCGTGGGACGTTTCATCTTTTACTGCTCGAAATATTTCCCTTCTCAG 600  
DB 736 AATGCGAGCAAGCGTGGGACGTTTCATCTTTTACTGCTCGAAATATTTCCCTTCTCAG 795  
QY 601 ATGCTGAGAAACGGGCAACAACTTTTACCTTTCAGTACACCTTTTGAGGAAGTGCCTTCCAC 660  
DB 796 ATGCTGAGAAACGGGCAACAACTTTTACCTTTCAGTACACCTTTTGAGGAAGTGCCTTCCAC 855  
QY 661 AGCAGTACGCGCACAGCCAGAGCTGGACCGGCTGATGAATCCTCTCATCCGACCAATAC 720  
DB 856 AGCAGTACGCGCACAGCCAGAGCTGGACCGGCTGATGAATCCTCTCATCCGACCAATAC 915  
QY 721 CTGTATTACCTCAACAGAACTCAAATCAGTCCGGAAGTCCCAAAACAAAGGACTTGTCTG 780  
DB 916 CUGTATTACCTGAAACAGAACTCAAATCAGTCCGGAAGTCCCAAAACAAAGGACTTGTCTG 975  
QY 781 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACTGGCTACCTGACACC 840  
DB 976 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACTGGCTACCTGACACC 1035  
QY 841 TGTATTCCGACAGCGCGTTCCTTAAACAAACAAACAGCAACAAACAGCAATTTTACC 900  
DB 1036 TGTATTCCGACAGCGCGTTCCTTAAACAAACAAACAGCAACAAACAGCAATTTTACC 1095  
QY 901 TGGACTGTGTCTCAAATATAACCTCAATGGCGGTGAATTCATCATCAACCTGGCACT 960  
DB 1096 TGGACTGTGTCTCAAATATAACCTCAATGGCGGTGAATTCATCATCAACCTGGCACT 1155  
QY 961 GCTATGGCTCTCACAAAGACGACGAAGTCTTTTCCCATGAGCGGTGTCAATGATT 1020  
DB 1156 GCTATGGCTCTCACAAAGACGACGAAGTCTTTTCCCATGAGCGGTGTCAATGATT 1215  
QY 1021 TTTGGAAAAGAGAGCGCCGAGCTTCAAAACATGCTTGGACAACTGCTATGATTACAGAC 1080  
DB 1216 TTTGGAAAAGAGAGCGCCGAGCTTCAAAACATGCTTGGACAACTGCTATGATTACAGAC 1275  
QY 1081 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCCGAAAGATTTGGGACCGTGGCAGTC 1140  
DB 1276 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCCGAAAGATTTGGGACCGTGGCAGTC 1335  
QY 1141 AATTTCCAGAGCAGCAGCAGACAGACCCCTGCGACCGGAGATGTGCTATGGGAGCATTA 1200  
DB 1336 AATTTCCAGAGCAGCAGCAGACAGACCCCTGCGACCGGAGATGTGCTATGGGAGCATTA 1395  
QY 1201 CTTGGCATGGTGTGGCAAGATAGAGAGCTGATCCTCGAGGTCCCATTTGGGCCAAAAT 1260  
DB 1396 CTTGGCATGGTGTGGCAAGATAGAGAGCTGATCCTCGAGGTCCCATTTGGGCCAAAAT 1455  
QY 1261 CCTCACACAGATGGACACTTTTCAACCGTCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 1320  
DB 1456 CCTCACACAGATGGACACTTTTCAACCGTCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 1515  
QY 1321 CCGCCCTCTCAGATCCTCATCAAAAACAGCCCTGTTTCTGCGAATCCTCGGCGGAGTTT 1380  
DB 1516 CCGCCCTCTCAGATCCTCATCAAAAACAGCCCTGTTTCTGCGAATCCTCGGCGGAGTTT 1575  
QY 1381 TCAGCTACAAAGTTTGGTTCATTCATCACCCTTCTCCAGAGACAGTGTGTGGAA 1440  
DB 1576 TCAGCTACAAAGTTTGGTTCATTCATCACCCTTCTCCAGAGACAGTGTGTGGAA 1635  
QY 1441 ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCTGGAAATCCCGAAGTGCAGTACACA 1500  
DB 1636 ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCTGGAAATCCCGAAGTGCAGTACACA 1695

QY 1501 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGAGCTTTATACT 1560  
Db |||||  
QY 1696 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGAGCTTTATACT 1755  
Db |||||  
QY 1561 GAGCCTCGCCCATTTGGCACCGTTACCTTACCGTCCCGCTGTA 1604  
Db 1756 GAGCCTCGCCCATTTGGCACCGTTACCTTACCGTCCCGCTGTA 1799

RESULT 3  
AAD00777  
ID AAD00777 standard; DNA; 2211 BP.  
XX  
AC AAD00777;  
XX  
DT 08-SEP-2000 (first entry)  
XX  
DE Adeno-associated virus serotype 1 capsid protein VP1 DNA.  
XX  
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;  
KW cap protein; recombinant viral vector; gene delivery; gene therapy;  
KW vaccine; transgene; VP1; ss.  
XX  
OS Adeno associated virus serotype 1.  
XX  
PH Location/Qualifiers  
FT 1..2211  
FT /tag= a  
FT /product= "VP1 protein"  
XX  
XX WO200028061-A2.  
XX  
XX 18-MAY-2000.  
XX  
XX 02-NOV-1999; 99WO-US025694.  
XX  
XX 05-NOV-1998; 98US-0107114P.  
XX  
XX (UYBE-) UNIV PENNSYLVANIA.  
XX  
XX Wilson JM, Xiao W;  
XX  
XX WPI: 2000-376571/32.  
XX P-P8DB; AAY71167.  
XX  
XX Novel adeno-associated virus serotype 1 polynucleotide useful for  
XX preparation of medicament for delivery of a transgene to a host.  
XX  
XX Claim 10; Page 83-87; 108pp; English.

XX The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA  
XX which is characterised by two inverted terminal repeats (ITR) and open  
XX reading frames for rep and capsid (cap) proteins. The rep reading frame  
XX encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap  
XX reading frame encodes three structural proteins, VP1, VP2 and VP3. The  
XX AAV-1 sequence or its fragments particularly ITRs, rep and cap coding  
XX regions, are useful in production of recombinant viral vectors for gene  
XX delivery. These vectors can be used as gene therapy vectors, vaccine  
XX vectors or antisense delivery vectors. The AAV-1 does not induce the  
XX formation of neutralising antibodies specific to any serotype of AAV  
XX hence is useful for transforming host cells, and in the preparation of a  
XX medicament for the delivery of transgene to a host. The present sequence  
XX is an AAV-1 DNA encoding a cap protein VP1 which is useful in the  
XX production of recombinant viral vector for gene delivery  
XX  
XX Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 U; 0 Other;  
XX  
XX Query Match 100.0%; Score 1604; DB 3; Length 2211;  
XX Best Local Similarity 100.0%; Pred. No. 0;  
XX Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1 ATGGCTTCAGCGGTGGCGCACCAATGGCAGACAAATAGGCGCGCACGAGTGGGT 60  
XX |||||

Db 607 ATGGCTTCAGCGGTGGCGCACCAATGGCAGACAAATAGGCGCGCACGAGTGGGT 666  
QY |||||  
QY 61 AATGCTCAGGAAATTTGGCATTTGCATATGGCTGGGGGACAGAGTCAATCAACACC 120  
Db |||||  
QY 667 AATGCTCAGGAAATTTGGCATTTGCATATGGCTGGGGGACAGAGTCAATCAACACC 726  
Db |||||  
QY 121 AGCACCGCACCTGGGCTTGGCCACCTTACAAATTAACACCTCTCAAGCAAAATCTCCAGT 180  
Db 727 AGCACCGCACCTGGGCTTGGCCACCTTACAAATTAACACCTCTCAAGCAAAATCTCCAGT 786  
QY |||||  
QY 181 GCTTCAACGGGGGCGCAGCAACGACCACTACTTTGGGCTACAGCACCCCTGGGGGTAT 240  
Db 787 GCTTCAACGGGGGCGCAGCAACGACCACTACTTTGGGCTACAGCACCCCTGGGGGTAT 846  
QY |||||  
QY 241 TTTGATTTCAACAGATTTCCACTGCGCATTTTTCACACGCTGAGCTGGCAGCGATCATCAAC 300  
Db 847 TTTGATTTCAACAGATTTCCACTGCGCATTTTTCACACGCTGAGCTGGCAGCGATCATCAAC 906  
QY |||||  
QY 301 AACAAATTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACCTCTTACCAACATCCAAGTCAAG 360  
Db 907 AACAAATTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACCTCTTCAACATCCAAGTCAAG 966  
QY |||||  
QY 361 GAGGTCAACGAGAAATGATGGGCTCACAAACCATCGCTAAATAACCTTTACAGCACGGTTCAA 420  
Db 967 GAGGTCAACGAGAAATGATGGGCTCACAAACCATCGCTAAATAACCTTTACAGCACGGTTCAA 1026  
QY |||||  
QY 421 GTCTTCTCGGACTGGAGTACAGGTTCCGTTACGTTCTCGGCTCTGGCGACAGAGGCTGC 480  
Db 1027 GTCTTCTCGGACTGGAGTACAGGTTCCGTTACGTTCTCGGCTCTGGCGACAGAGGCTGC 1086  
QY |||||  
QY 481 CTCCCTCCGTTCCGGCGGAGCTGTTCAATGATTCGGCAATACGGCTACTGACGCTCAAC 540  
Db 1087 CTCCCTCCGTTCCGGCGGAGCTGTTCAATGATTCGGCAATACGGCTACTGACGCTCAAC 1146  
QY |||||  
QY 541 AATGGCAGCCAAAGCGTGGGAGCTTCACTCTTTTACTGCTCGAATAATTTCCCTTCTCAG 600  
Db 1147 AATGGCAGCCAAAGCGTGGGAGCTTCACTCTTTTACTGCTCGAATAATTTCCCTTCTCAG 1206  
QY |||||  
QY 601 ATGCTGAGAACGGGCAACAACCTTTACCTTACGTTACCTTTGAGGAAGTGGCTTTCCAC 660  
Db 1207 ATGCTGAGAACGGGCAACAACCTTTACCTTACGTTACCTTTGAGGAAGTGGCTTTCCAC 1266  
QY |||||  
QY 661 AGCAGCTACGGCGCACAGCCAGAGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720  
Db 1267 AGCAGCTACGGCGCACAGCCAGAGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 1326  
QY |||||  
QY 721 CTGTATTACCTGAAACAGAACTCAAATCAGTCCGGAAGTGGCCAAACAAAGGACTTGTCTG 780  
Db 1327 CTGTATTACCTGAAACAGAACTCAAATCAGTCCGGAAGTGGCCAAACAAAGGACTTGTCTG 1386  
QY |||||  
QY 781 TTTAGCCGTGGTCTCCAGCTGGCATGTTCTGTTAGCCCAAAAACCTGCTACTCGACCC 840  
Db 1387 TTTAGCCGTGGTCTCCAGCTGGCATGTTCTGTTAGCCCAAAAACCTGCTACTCGACCC 1446  
QY |||||  
QY 841 TGTATTACGGCAGCGCGGTTTCTAAAACAAAACAGACAAACAACAGCAATTTTACC 900  
Db 1447 TGTATTACGGCAGCGCGGTTTCTAAAACAAAACAGACAAACAACAGCAATTTTACC 1506  
QY |||||  
QY 901 TGGACTGTGCTTCAAAAATATAAATCTCAATGGCGTGAATCCATCATCAACCTCGCACT 960  
Db 1507 TGGACTGTGCTTCAAAAATATAAATCTCAATGGCGTGAATCCATCATCAACCTCGCACT 1566  
QY |||||  
QY 961 GCTATGGCTTACACAAAAGACGAAAGATTTCTTCCCATGAGCGGTGTCTATGATT 1020  
Db 1567 GCTATGGCTTACACAAAAGACGAAAGATTTCTTCCCATGAGCGGTGTCTATGATT 1626  
QY |||||  
QY 1021 TTTTGAAAAGAGAGCGCGGAGCTTCAAACTGCTATTTGGCAATGTCTATGATTACAGAC 1080  
Db 1627 TTTTGAAAAGAGAGCGCGGAGCTTCAAACTGCTATTTGGCAATGTCTATGATTACAGAC 1686  
QY |||||  
QY 1081 GAAGAGGAAATTTAAAGCCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTTC 1140  
Db 1687 GAAGAGGAAATTTAAAGCCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTTC 1746







Db 1387 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTCCAGCCAAAACTGGCTACCTGCAGCC 1446  
Qy 841 TGTATTCCGACAGCGGCTTTCTAAACAAAAAGACAAACACAGCAATTTTACC 900  
Db 1447 TGTATTCCGACAGCGGCTTTCTAAACAAAAAGACAAACACAGCAATTTTACC 1506  
Qy 901 TGGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 960  
Db 1507 TGGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 1566  
Qy 961 GCTATGGCTCTACAAAAGACGACGAAGAACAAGTTCTTTCCATGAGCGGTGTATGATT 1020  
Db 1567 GCTATGGCTCTACAAAAGACGACGAAGAACAAGTTCTTTCCATGAGCGGTGTATGATT 1626  
Qy 1021 TTTGGAAAAGAGCGCCGGAGCTTCAACACTGTGATTTGGACAATGTCTATGATTCAGAC 1080  
Db 1627 TTTGGAAAAGAGCGCCGGAGCTTCAACACTGTGATTTGGACAATGTCTATGATTCAGAC 1686  
Qy 1081 GAAGAGGAATTTAAAGCCACTAACCTGTGGCCACCAGAAAGATTTGGGACCGTGGCAGTC 1140  
Db 1687 GAAGAGGAATTTAAAGCCACTAACCTGTGGCCACCAGAAAGATTTGGGACCGTGGCAGTC 1746  
Qy 1141 AATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCTATGAGGACATTA 1200  
Db 1747 AATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCTATGAGGACATTA 1806  
Qy 1201 CTTGGCATGGTGGCAAGATAGAGACGTGTACTTCAGGGTCCCATTTTGGGCCAAAATT 1260  
Db 1807 CTTGGCATGGTGGCAAGATAGAGACGTGTACTTCAGGGTCCCATTTTGGGCCAAAATT 1866  
Qy 1261 CTTCAACAGATGGACACTTTTCAACCGTCTCTCTTTATGGCGGCTTTGGACTCAAGAAC 1320  
Db 1867 CTTCAACAGATGGACACTTTTCAACCGTCTCTCTTTATGGCGGCTTTGGACTCAAGAAC 1926  
Qy 1321 CCGCTCTCTCAGATCCTCATCAAAAACACGCTGTCTTCTGCGAATCTCTCGGCGGAGTTT 1380  
Db 1927 CCGCTCTCTCAGATCCTCATCAAAAACACGCTGTCTTCTGCGAATCTCTCGGCGGAGTTT 1986  
Qy 1381 TCAGCTACAAAGTTTGCTTCATTCACCCATCTACCCATCTCCAGGACAAAGTGTGGAA 1440  
Db 1987 TCAGCTACAAAGTTTGCTTCATTCACCCATCTCCAGGACAAAGTGTGTGGAA 2046  
Qy 1441 ATTGAATGGAGCTGCAGAAAAGAAACAGACGCTGGAATCCCGAAGTGCAATACACA 1500  
Db 2047 ATTGAATGGAGCTGCAGAAAAGAAACAGACGCTGGAATCCCGAAGTGCAATACACA 2106  
Qy 1501 TCCAAATTATGCAAAATCTGCCAAGTTGATTTTACTGTGGACAAACATGGACTTTATCT 1560  
Db 2107 TCCAAATTATGCAAAATCTGCCAAGTTGATTTTACTGTGGACAAACATGGACTTTATCT 2166  
Qy 1561 GAGCCTGGCCCAATTGGCACCGGTTACCTTACCGTCCCTCTGA 1604  
Db 2167 GAGCCTGGCCCAATTGGCACCGGTTACCTTACCGTCCCTCTGA 2210

## RESULT 5

AD226929

ID AD226929 standard; DNA; 2211 BP.

XX

AC AD226929;

XX

DT 30-JUN-2005 (first entry)

XX

DE Adeno-associated virus DNA SEQ ID NO 79.

XX

KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;  
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;  
KW bacterial infection; cancer; ulcerative colitis; anti-rheumatic;  
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;  
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;  
KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.  
XX  
OS Adeno-associated virus.

XX WO200503321-A2.  
PN 14-APR-2005.  
XX 30-SEP-2004; 2004WO-US028817.  
PF 30-SEP-2003; 2003US-0508226P.  
PR 29-APR-2004; 2004US-0566546P.  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;  
PI WPI; 2005-285437/29.  
XX  
DR New adeno-associated virus (AAV) clade comprising at least three AAV  
XX members, useful for preventing and/or treating arthritis, multiple  
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial  
PT infection and cancer.  
XX  
PS Disclosure; SEQ ID NO 79; 569pp; English.  
XX  
CC The invention relates to an adeno-associated virus (AAV) clade comprising  
CC at least three AAV members, where each member of the AAV clade is  
CC phylogenetically related as determined using a Neighbor-Joining heuristic  
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson  
CC correction distance measurement of no more than 0.05. The methods and  
CC compositions of the present invention are useful for the prevention  
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,  
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's  
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative  
CC colitis. The present sequence represents an adeno-associated virus DNA.  
XX  
SQ Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1604; DB 14; Length 2211;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGCTTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGTGGGT 60  
Db 607 ATGGCTTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGTGGGT 666  
Qy 61 AATGCTCAGGAAATTTGGCATTTGCATTCACATGGCTGGGCGCAGAGTCAACACC 120  
Db 667 AATGCTCAGGAAATTTGGCATTTGCATTCACATGGCTGGGCGCAGAGTCAACACC 726  
Qy 121 AGCACCCGACCTGGGCGCTTGGCCACCTACCAATACCACTTCAAGCAAAATCTCCAGT 180  
Db 727 AGCACCCGACCTGGGCGCTTGGCCACCTACCAATACCACTTCAAGCAAAATCTCCAGT 786  
Qy 181 GCTTCAACGGGGCGCCAGCAACGACCACTACTTTGGGCTACAGCACCCCTGGGGGTAT 240  
Db 787 GCTTCAACGGGGCGCCAGCAACGACCACTACTTTGGGCTACAGCACCCCTGGGGGTAT 846  
Qy 241 TTTGATTTCAACAGATTTCCATGCGCACCTTTTCCACAGCTGACTGCGAGGACTCATCAAC 300  
Db 847 TTTGATTTCAACAGATTTCCATGCGCACCTTTTCCACAGCTGACTGCGAGGACTCATCAAC 906  
Qy 301 AACAAATGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCAAGTCAAG 360  
Db 907 AACAAATGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCAAGTCAAG 966  
Qy 361 GAGGTCAACGCAATGATGGCGTCAACACCATCGCTAATAACCTTACCAGCACGGTTCAA 420  
Db 967 GAGGTCAACGCAATGATGGCGTCAACACCATCGCTAATAACCTTACCAGCACGGTTCAA 1026  
Qy 421 GTCTTCTCGGACTCGGAGTACAGGCTTCGTACGCTCTCGGCTCTGCGCACCGAGGCTGC 480  
Db 1027 GTCTTCTCGGACTCGGAGTACAGGCTTCGTACGCTCTCGGCTCTGCGCACCGAGGCTGC 1086  
Qy 481 CTCCCTCGGTTCCCGCGGAGCGTGTTCATGATTCCGCAATACGGCTACTGTACGCTCAAC 540

Db 1087 |||||CTCCCTCCGTTCCCGCGACGTTTCATGATTCGCGAATACGGCTACCTGACGCTCAAC 1146  
QY 541 AATGGCAGCAAGCGCTGGAGCTTTCATCTCTTTACTGCTGGAATATTTCTCTCTCAG 600  
Db 1147 AATGGCAGCAAGCGCTGGAGCTTTCATCTCTTTACTGCTGGAATATTTCTCTCTCAG 1206  
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTCCAC 660  
Db 1207 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTCCAC 1266  
QY 661 AGCAGCTACGGCAGCAGCGCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720  
Db 1267 AGCAGCTACGGCAGCAGCGCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 1326  
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAAGAACTTGGTG 780  
Db 1327 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAAGAACTTGGTG 1386  
QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAACTGGCTACCTGGACCC 840  
Db 1387 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAACTGGCTACCTGGACCC 1446  
QY 841 TGTATTTCGCAGCAGCGCTTTCTAATAAATAAAGCAGCAAAAGCAATTTTACC 900  
Db 1447 TGTATTTCGCAGCAGCGCTTTCTAATAAATAAAGCAGCAAAAGCAATTTTACC 1506  
QY 901 TGGACTGGTCTCAAAATATACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 960  
Db 1507 TGGACTGGTCTCAAAATATACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 1566  
QY 961 GCTATGGCTTCACACAAAGCAGCAGCAAGTCTTTCCATGAGCGGTGTCATGATT 1020  
Db 1567 GCTATGGCTTCACACAAAGCAGCAGCAAGTCTTTCCATGAGCGGTGTCATGATT 1626  
QY 1021 TTTGAAAAAGAGCAGCGCGAGCTTCAAAACATGTGCATTTGGACAATGTCATGATTAAGAC 1080  
Db 1627 TTTGAAAAAGAGCAGCGCGAGCTTCAAAACATGTGCATTTGGACAATGTCATGATTAAGAC 1686  
QY 1081 GAAGAGGAATTAAGCCCTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140  
Db 1687 GAAGAGGAATTAAGCCCTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1746  
QY 1141 AATTTCCAGAGCAGCAGACAGACCTTCGACCGAGATGTGCATGCTATGGGAGCATTA 1200  
Db 1747 AATTTCCAGAGCAGCAGACAGACCTTCGACCGAGATGTGCATGCTATGGGAGCATTA 1806  
QY 1201 CCTGGCATGGTGGCAAGATAGAGACGTGTACCTGACGGTCCCAATTTGGGCCAAAAT 1260  
Db 1807 CCTGGCATGGTGGCAAGATAGAGACGTGTACCTGACGGTCCCAATTTGGGCCAAAAT 1866  
QY 1261 CCTCACACAGATGGACATTTTCACCCGTCTCCTTTATGGCGGCTTTGGACTCAAGAAC 1320  
Db 1867 CCTCACACAGATGGACATTTTCACCCGTCTCCTTTATGGCGGCTTTGGACTCAAGAAC 1926  
QY 1321 CGCCCTCTCAGATCCTCATCAAAAACAGCCCTGTTCTTCGCAATCTCCGGCGGAGTTT 1380  
Db 1927 CGCCCTCTCAGATCCTCATCAAAAACAGCCCTGTTCTTCGCAATCTCCGGCGGAGTTT 1986  
QY 1381 TCAGCTACAAAGTTTGCTTCATTCATCACCCAACTACTCCAAGGACAAGTGAAGTGGAA 1440  
Db 1987 TCAGCTACAAAGTTTGCTTCATTCATCACCCAACTACTCCAAGGACAAGTGAAGTGGAA 2046  
QY 1441 ATTGAATGGAGCTGCAGAAAGAAAACAGCAACCGCTGGAATCCGAAGTGCAGTACACA 1500  
Db 2047 ATTGAATGGAGCTGCAGAAAGAAAACAGCAACCGCTGGAATCCGAAGTGCAGTACACA 2106  
QY 1501 TCCAAATATGCAAAATCTGCAACGTTGATTTTACTGTGGACAACAAATGGAATTTTACT 1560  
Db 2107 TCCAAATATGCAAAATCTGCAACGTTGATTTTACTGTGGACAACAAATGGAATTTTACT 2166  
QY 1561 GAGCCTCGCCCATTTGGCACCCGTTTACCTTTACCCGCTCCCTGTA 1604  
|||||

Db 2167 GAGCCTCGCCCATTTGGCACCGTTACCTTACCCGCTCCCTGTGA 2210  
RESULT 6  
ADW39398  
ID ADW39398 standard; DNA; 4347 BP.  
XX  
AC ADW39398;  
XX  
DT 24-MAR-2005 (first entry)  
XX  
DE Adeno-associated virus 2 (AAV2) rep-AAV1 cap fusion DNA - SEQ ID 1.  
XX  
KW genetically engineered microorganism; vector; rep; cap; ds.  
XX  
OS Adeno-associated virus 2.  
OS Chimeric.  
OS Unidentified.  
XX  
PN CN1461805-A.  
XX  
PD 17-DEC-2003.  
XX  
PF 27-MAY-2002; 2002CN-00117965.  
XX  
PR 27-MAY-2002; 2002CN-00117965.  
XX  
PA (BENV-) BENYUAN ZHENGYANG GENETIC TECHNOLOGY CO.  
XX  
PI Wu X, Cao H, Dong X;  
XX  
DR WPI; 2004-248658/24.  
XX  
PT Structure of a recombinant herpes simplex virus and its use.  
XX  
PS Claim 1; SEQ ID NO 1; 102pp; Chinese.  
XX  
CC The invention relates to a novel recombinant herpes simplex virus (rHSVS) whose genome is respectively inserted by the Adeno-associated virus (AAV) rep gene and cap gene. The AAV genes may be derived from viruses AAV1, AAV2, AAV3, AAV4, AAV5 and AAV6. The methods of the invention may be useful for generating recombinant vectors with high efficiency and no need of reconfiguration. The current sequence is that of the Adeno-associated virus 2 (AAV2) rep-Adeno-associated virus 1 (AAV1) cap fusion DNA (SEQ ID 1) of the invention.  
XX  
SQ Sequence 4347 BP; 1106 A; 1186 C; 1142 G; 913 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1604; DB 13; Length 4347;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGCAATAGAGAGCGCCGACGAGTGGGT 60  
Db 2627 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGCAATAGAGAGCGCCGACGAGTGGGT 2686  
QY 61 AATGCTCAGGAAATTTGGCATTCGATTCACATGGCTGGCGGACAGAGTCTACACC 120  
Db 2687 AATGCTCAGGAAATTTGGCATTCGATTCACATGGCTGGCGGACAGAGTCTACACC 2746  
QY 121 AGCACCCGACCTGGGCGCTTGGCCACCTACATAAACCACTCTTACAGCAATCTCCAGT 180  
Db 2747 AGCACCCGACCTGGGCGCTTGGCCACCTACATAAACCACTCTTACAGCAATCTCCAGT 2806  
QY 181 GCTTCAACGGGGCGCAGCAACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT 240  
Db 2807 GCTTCAACGGGGCGCAGCAACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT 2866  
QY 241 TTTGATTTCAACAGAGTTCCACTGGCACTTTTCCACACGTGACTGGCAGCACTCATCAAC 300  
Db 2867 TTTGATTTCAACAGAGTTCCACTGGCACTTTTCCACACGTGACTGGCAGCACTCATCAAC 2926  
QY 301 AACAAATTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 360

Db 2927 AACAAATGGGATTCGGGCCAAGAGACTCAACTCTTCAAACTCTTCAACATCCAGTCAAG 2986  
QY 361 GAGGTACAGCAAGATGATGGCTCAACCAATCGCTAATAACCTTTACCAAGCAGCGTTCAA 420  
Db 2987 GAGGTACAGCAAGATGATGGCTCAACCAATCGCTAATAACCTTTACCAAGCAGCGTTCAA 3046  
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGCTCTCGGCTCTGGCGCACCGGCTGC 480  
Db 3047 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGCTCTCGGCTCTGGCGCACCGGCTGC 3106  
QY 481 CTCCCTCGGTTCCCGCGAGCGTGTTCATGATTCGCAATACGCTACCTGACGCTCAAC 540  
Db 3107 CTCCCTCGGTTCCCGCGAGCGTGTTCATGATTCGCAATACGCTACCTGACGCTCAAC 3166  
QY 541 AATGGCAGCAAGCCGTGGAGCTTCATCTTTTACTGCTCGAATAATTTCCCTTCTCAG 600  
Db 3167 AATGGCAGCAAGCCGTGGAGCTTCATCTTTTACTGCTCGAATAATTTCCCTTCTCAG 3226  
QY 601 ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCAC 660  
Db 3227 ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCAC 3286  
QY 661 AGCAGCTACGCGACAGCAGAGCTTGGACCGCTGATGAATCCTCTCATCGACCAATAC 720  
Db 3287 AGCAGCTACGCGACAGCAGAGCTTGGACCGCTGATGAATCCTCTCATCGACCAATAC 3346  
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAAAGGACTTGCTG 780  
Db 3347 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAAAGGACTTGCTG 3406  
QY 781 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAAACCTGGCTACTGGACCC 840  
Db 3407 TTTAGCCGTGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAAACCTGGCTACTGGACCC 3466  
QY 841 TGTATTCGCGACGCGCGTCTTCTAAACAAACAAACAGACAAACAAACAGCAATTTTACC 900  
Db 3467 TGTATTCGCGACGCGCGTCTTCTAAACAAACAAACAGACAAACAAACAGCAATTTTACC 3526  
QY 901 TGACCTGGTGTCTTAAATATATACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 960  
Db 3527 TGACCTGGTGTCTTAAATATATACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 3586  
QY 961 GCTATGGCTTCACAAAAGACGAGCAAGATGTTCTTTCCCATGAGCGGTGCATGATT 1020  
Db 3587 GCTATGGCTTCACAAAAGACGAGCAAGATGTTCTTTCCCATGAGCGGTGCATGATT 3646  
QY 1021 TTTGGAAAAGAGCGCGGAGCTTCAACACTGCAATTCGACAAATGTCATGATTACAGAC 1080  
Db 3647 TTTGGAAAAGAGCGCGGAGCTTCAACACTGCAATTCGACAAATGTCATGATTACAGAC 3706  
QY 1081 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140  
Db 3707 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 3766  
QY 1141 AATTTCCAGACGACGACGACGACCTTCGACCGGAGATGTGCATGCTATGGGAGCATTA 1200  
Db 3767 AATTTCCAGACGACGACGACGACCTTCGACCGGAGATGTGCATGCTATGGGAGCATTA 3826  
QY 1201 CTTGGCATGCTGTGGCAAGATAGAGAGTGTACCTGCGAGTCCCATTTTGGGCCAAATTT 1260  
Db 3827 CTTGGCATGCTGTGGCAAGATAGAGAGTGTACCTGCGAGTCCCATTTTGGGCCAAATTT 3886  
QY 1261 CCTCACAGATGGACATTTTCAACCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1320  
Db 3887 CCTCACAGATGGACATTTTCAACCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 3946  
QY 1321 CGGCTCTCAGATCTCTCATCAAAAACAGCCTGTTCTCGGAACTCTCGGCGGAGTTT 1380  
Db 3947 CGGCTCTCAGATCTCTCATCAAAAACAGCCTGTTCTCTCGGAATCTCTCGGCGGAGTTT 4006  
QY 1381 TCAGCTACAAAGTTTCTTCATTCATCCCAATCTCCACAGGACAAAGTGAAGTGGAA 1440

Db 4007 TCAGTCAAAAGTTTGCTTTCATTCACCACTATCCACAGGACAAGTGAAGTGGAA 4066  
QY 1441 ATTGAATGGGAGCTGCGAGAAAAGAAAAGCAAGCGCTGGAATCCGAAAGTGCAGTACACA 1500  
Db 4067 ATTGAATGGGAGCTGCGAGAAAAGAAAAGCAAGCGCTGGAATCCGAAAGTGCAGTACACA 4126  
QY 1501 TCCAAATTATGCAGAAATCTGCCAAGCTTGTGATTTTACTGTGGACACAAATGGGACTTTATCT 1560  
Db 4127 TCCAAATTATGCAGAAATCTGCCAAGCTTGTGATTTTACTGTGGACACAAATGGGACTTTATCT 4186  
QY 1561 GAGCTCTGCCCTTATGGCACCCCTTACCTTACCCGTCCCTGTGA 1604  
Db 4187 GAGCTCTGCCCTTATGGCACCCCTTACCTTACCCGTCCCTGTGA 4230  
RESULT 7  
ADZ46594  
ID ADZ46594 standard; DNA; 4347 BP.  
XX ADZ46594;  
AC ADZ46594;  
XX 30-JUN-2005 (first entry)  
XX HSV-AAV sequence rep2cap1 for AAV/HSV vectors.  
XX Genetic engineering; gene therapy; vector; cancer; neoplasm; cytostatic;  
KW respiratory disease; respiratory-gen.; neurological disease;  
KW neuroprotective; ds.  
XX Adeno-associated virus.  
OS Human herpesvirus 1.  
OS WO2005035743-A1.  
PN 21-APR-2005.  
PD 15-OCT-2003; 2003WO-CN000861.  
PF 15-OCT-2003; 2003WO-CN000861.  
PR (AGTC-) AGTC GENE TECHNOLOGY CO LTD.  
XX Wu X, Cao H, Dong X;  
PI WPI; 2005-296279/30.  
DR Large-scale production, isolation and purification of serotype adeno-associated virus vectors by infecting cells with recombinant herpes simplex virus vectors, for use in gene therapy of e.g. cancer and respiratory diseases.  
PT Claim 1; SEQ ID NO 1; 79pp; Chinese.  
PS The invention relates to a recombinant herpes simplex virus (HSV) characterized in that its genome is inserted with a DNA sequence selected from nucleotide sequences of ADZ46594-ADZ46598 or their homologous sequences. Also included are preparing the recombinant HSV (comprising the construction of DNA fragments that contains sequences ADZ46594-ADZ46598, and respectively inserting these 5 DNA fragments into genome of HSV by applying genetic engineering to give the recombinant HSV), large-scale production of 5 serotype recombinant adeno-associated viruses (AAV) 1, 3, 4, 5, and 6 (by preparing the recombinant HSV, establishing a one-strain vector cell i.e. recombinant AAV vector cell-line, infecting the recombinant AAV under the infective effect of the 5 recombinant HSV in the vector cell-line, separating and purifying the 5 serotype recombinant AAV after lysing the AAV-containing cells and culture liquor to give a crude lysate and further purification of rAAV by density-gradient centrifugation or affinity chromatography) and a recombinant vector plasmid pSNV-NX (containing ITR at both ends of AAV-1, AAV-3, AAV-4, AAV-5 or AAV-6 genome, with immediate-early enhancer and promoter of cytomegalovirus, and a polyclonal site and a polyA signal, respectively, between the ITR (inverted terminal repeat), and neomycin-resistance gene-

CC expressing cassette at outer edge of ITR). The DNA sequence is inserted  
CC into XbaI site in UL2 or UL4 gene in HSV genome. The DNA sequence of  
CC sequence AD246594-AD246598 can also be inserted into other non-essential  
CC gene regions in HSV genome. The recombinant HSV is optionally inserted  
CC with other DNA sequences homologous with the already-specified fragments.  
CC The vectors are for use in gene therapy of e.g. cancer, respiratory  
CC diseases and neural diseases. The virus vectors are safe, with long  
CC expression time and wide-spectrum of cell infection, even non-cleaved  
CC cells and reverse axonal conduction through the incorporated HSV vectors,  
CC and high transfer efficiency. The present sequence is the HSV-AAV  
CC sequence rep2cap1 for the AAV/HSV vectors of the invention.  
XX  
SQ Sequence 4347 BP; 1106 A; 1186 C; 1142 G; 913 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1604; DB 14; Length 4347;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTTCAGGGCGTGGCGCACCACCAATGGCAGACAATAACGAAGGCGCCGACGGAGTGGGT 60  
Db 2627 ATGGCTTCAGGGCGTGGCGCACCACCAATGGCAGACAATAACGAAGGCGCCGACGGAGTGGGT 2686  
QY 61 AATGCCCTCAGGAAATTTGGCATTTCCGATTTCCATGCTGGCGCAGAGTTCATCACCAC 120  
Db 2687 AATGCCCTCAGGAAATTTGGCATTTCCGATTTCCATGCTGGCGCAGAGTTCATCACCAC 2746  
QY 121 AGCACCCGACCTGGGCGCTTGGCCACCTACATAACCACTCTACAAAGCAAAATCTCCAGT 180  
Db 2747 AGCACCCGACCTGGGCGCTTGGCCACCTACATAACCACTCTACAAAGCAAAATCTCCAGT 2806  
QY 181 GCTTCAACGGGGGCCAGCAACCAACCACTACTCTCGGCTACAGCAACCCCTGGGGGTAT 240  
Db 2807 GCTTCAACGGGGGCCAGCAACCAACCACTACTCTCGGCTACAGCAACCCCTGGGGGTAT 2866  
QY 241 TTTGATTTCAACAGATTCCA CTGCCA CTTTTACCAAGTGGA CTGGCAGGACTCATCAAC 300  
Db 2867 TTTGATTTCAACAGATTCCA CTGCCA CTTTTACCAAGTGGA CTGGCAGGACTCATCAAC 2926  
QY 301 AACAATTTGGGATTTCCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 360  
Db 2927 AACAATTTGGGATTTCCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 2986  
QY 361 GAGGTCAACGAGAAATGATGGCGTCAACCAATCGCTTAATAACCTTACAGCAGCGTTCAA 420  
Db 2987 GAGGTCAACGAGAAATGATGGCGTCAACCAATCGCTTAATAACCTTACAGCAGCGTTCAA 3046  
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGTCCTGGCTCTGGGACACAGGCGCTGC 480  
Db 3047 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGTCCTGGCTCTGGGACACAGGCGCTGC 3106  
QY 481 CTCCCTCCGTTCCGGGGGACGTTTCATGATTCGGCAATACGGCTACCTGACGCTCAAC 540  
Db 3107 CTCCCTCCGTTCCGGGGGACGTTTCATGATTCGGCAATACGGCTACCTGACGCTCAAC 3166  
QY 541 AATGGCAGCAACGCGTGGGACGTTTCATCTCTTTTACTGCTTGAATAATTTCCCTTCTCAG 600  
Db 3167 AATGGCAGCAACGCGTGGGACGTTTCATCTTTTACTGCTTGAATAATTTCCCTTCTCAG 3226  
QY 601 ATGCTGAGAACGGGCAACAACTTTACCTTCAGTACACCTTTGAGGAAGTGCCTTTCCAC 660  
Db 3227 ATGCTGAGAACGGGCAACAACTTTACCTTCAGTACACCTTTGAGGAAGTGCCTTTCCAC 3286  
QY 661 AGCAGTACGGCAGCAGCAGAGCGCTGGACCGGCTGATGATCTCTCTACGACCAATAC 720  
Db 3287 AGCAGTACGGCAGCAGCAGAGCGCTGGACCGGCTGATGATCTCTCTACGACCAATAC 3346  
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAGAGACTTGTCTG 780  
Db 3347 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAGAGACTTGTCTG 3406  
QY 781 TTTAGCGTGGGTCTCCAGCTGGCATGTCTGTTTCCAGCCCAAAACTGGCTTACCTGGACCC 840  
Db 3407 TTTAGCGTGGGTCTCCAGCTGGCATGTCTGTTTCCAGCCCAAAACTGGCTTACCTGGACCC 3466

QY 841 TGTATCGCAGCAGCGCGTCTTCTAAACAAACAAACAGACAAACAGCAATTTTACC 900  
Db 3467 TGTATCGCAGCAGCGCGTCTTCTAAACAAACAAACAGACAAACAGCAATTTTACC 3526  
QY 901 TGGACTGTGCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 960  
Db 3527 TGGACTGTGCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 3586  
QY 961 GCTATGGCTCTACACAAAGAGACGACCAAGATTTCTTTCCATGAGGGGTGTCATGATT 1020  
Db 3587 GCTATGGCTCTACACAAAGAGACGACCAAGATTTCTTTCCATGAGGGGTGTCATGATT 3646  
QY 1021 TTTGAAAGAGAGCGCGGAGCTTCAACACTGCAATTTGGACAATGTCATGATTCAGAC 1080  
Db 3647 TTTGAAAGAGAGCGCGGAGCTTCAACACTGCAATTTGGACAATGTCATGATTCAGAC 3706  
QY 1081 GAGAGGAAATTAAGGCACTTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140  
Db 3707 GAGAGGAAATTAAGGCACTTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 3766  
QY 1141 AATTTCCAGAGCAGCAGCAGCAGCCTGCGACCGGAGATGTGCTATGGGAGCATTA 1200  
Db 3767 AATTTCCAGAGCAGCAGCAGCAGCCTGCGACCGGAGATGTGCTATGGGAGCATTA 3826  
QY 1201 CTTGCAATGTTGGCAAGATAGACAGCTGTACCTGCAAGGTTCCCATTTTGGGCCAAAATT 1260  
Db 3827 CTTGCAATGTTGGCAAGATAGACAGCTGTACCTGCAAGGTTCCCATTTTGGGCCAAAATT 3886  
QY 1261 CTTCAACAGATGGACATTTTCCCGCTCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 1320  
Db 3887 CTTCAACAGATGGACATTTTCCCGCTCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 3946  
QY 1321 CGGCTCTCTCAGATCTCATCAAAACAGCGCTGTTCTTGGCAATCTCTCGCGGAGTTT 1380  
Db 3947 CGGCTCTCTCAGATCTCATCAAAACAGCGCTGTTCTTGGCAATCTCTCGCGGAGTTT 4006  
QY 1381 TCAGCTACAAAGTTTGTCTTCAATCATCCCAATATCTCCAGGACAAAGTGTGGA 1440  
Db 4007 TCAGCTACAAAGTTTGTCTTCAATCATCCCAATATCTCCAGGACAAAGTGTGGA 4066  
QY 1441 ATTGAATGGAGCTGCGAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500  
Db 4067 ATTGAATGGAGCTGCGAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 4126  
QY 1501 TCCAAATATGCAAAATCTGCCAAAGTGTGATTTTACTGTGCAACAAATGGACTTTTACT 1560  
Db 4127 TCCAAATATGCAAAATCTGCCAAAGTGTGATTTTACTGTGCAACAAATGGACTTTTACT 4186  
QY 1561 GAGCCTCGCCCATTTGGCACCCTTACCTTACCCCGTCCCTGTA 1604  
Db 4187 GAGCCTCGCCCATTTGGCACCCTTACCTTACCCCGTCCCTGTA 4230

RESULT 8

AAD00772

ID AAD00772 standard; DNA; 4718 BP.

XX AAD00772;

XX 08-SEP-2000 (first entry)

XX Adeno-associated virus serotype 1 DNA.

XX Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;  
KW cap protein; recombinant viral vector; gene delivery; gene therapy;  
KW vaccine; transgene; ss.  
XX  
XX Adeno associated virus serotype 1.

Key Location/Qualifiers

FT repeat\_unit 1..143

FT /\*tag= a

```
FT /label= 5' ITR
FT /note= "Inverted terminal repeat which is capable of
FT forming T-shaped hairpin structure"
FT 89..110
FT /tag= b
FT /bound_moiety= "Rep protein"
FT 124..125
FT /tag= c
FT /note= "Terminal resolute site (TRS)"
FT 219..226
FT /tag= d
FT /bound_moiety= "USF"
FT 236..299
FT /tag= e
FT /label= P5_promoter
FT 237..245
FT /tag= f
FT /bound_moiety= "YY1 factor"
FT 270..275
FT /tag= g
FT /label= P5_TATA-Box
FT 299..306
FT /tag= h
FT /note= "YY1/p5 RNA"
FT 335..2272
FT /tag= j
FT /product= "Rep 68"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT /note= "The coding region is interrupted by intron"
FT 335..2206
FT /tag= i
FT /product= "Rep 78"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT 857..862
FT /tag= l
FT /label= P19_TATA_Box
FT 882..883
FT /tag= m
FT /note= "P19 RNA"
FT 1007..2272
FT /tag= o
FT /product= "Rep 40"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT /note= "The coding region is interrupted by intron"
FT 1007..2206
FT /tag= n
FT /product= "Rep 52"
FT /function= "regulates replication and integration of AAV
FT DNA into host cell's chromosome"
FT 1836..1841
FT /tag= p
FT /label= P40_TATA-BOX
FT 1875..1876
FT /tag= q
FT /note= "P40 RNA"
FT 1924..2220
FT /tag= k
FT /note= "This region interrupts the coding sequence of Rep
FT 68 and Rep 40"
FT 2223..4433
FT /tag= r
FT /product= "VP1 protein"
FT /note= "Capsid protein"
FT 2634..4433
FT /tag= s
FT /product= "VP2 protein"
FT /note= "Capsid protein"
FT /partial
FT 2829..4433
```

```
FT /tag= t
FT /product= "VP3 protein"
FT /note= "Capsid protein"
FT 4447..4452
FT /tag= u
FT 4576..4718
FT /tag= v
FT /label= 3' ITR
FT /note= "Inverted terminal repeat which is capable of
FT forming T-shaped hairpin structure"
FT
XX WO200028061-A2.
XX
XX 18-MAY-2000.
XX
XX 02-NOV-1999; 99WO-US025694.
XX
XX 05-NOV-1998; 98US-0107114P.
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Wilson JM, Xiao W;
XX
XX WPI; 2000-376571/32.
XX P-PSDB; AAY71161, AAY71164, AAY71165, AAY71166, AAY71167, AAY71168,
XX AAY71169.
XX
XX Novel adeno-associated virus serotype 1 polynucleotide useful for
XX preparation of medicament for delivery of a transgene to a host.
XX
XX Claim 1; Fig 1; 108pp; English.
XX
XX The present sequence is an adeno-associated virus serotype 1 (AAV-1) DNA
XX characterised by two inverted terminal repeats (ITR) and open reading
XX frames for rep and capsid (cap) proteins. The rep reading frame encodes
XX four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap reading
XX frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1
XX sequence or its fragments particularly ITRs, rep and cap coding regions,
XX are useful in production of recombinant viral vectors for gene delivery.
XX These vectors can be used as gene therapy vectors, vaccine vectors or
XX antisense delivery vectors. The AAV-1 does not induce the formation of
XX neutralising antibodies specific to any serotype of AAV hence is useful
XX for transforming host cells, and in the preparation of a medicament for
XX the delivery of transgene to a host
XX
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 1604; DB 3; Length 4718;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ATGGCTTCAGCGGTGGCGCCACCAATGGCAGACAATAACGAAGGCCGACGGAGTGGGT 60
Db 2829 ATGGCTTCAGCGGTGGCGCCACCAATGGCAGACAATAACGAAGGCCGACGGAGTGGGT 2888
QY 61 AATGCCTCAGGAATTTGGCATTGCCATTCGATCCCATGGCTGGGCGCAGAGTCATCACCACC 120
Db 2889 AATGCCTCAGGAATTTGGCATTGCCATTCGATCCCATGGCTGGGCGCAGAGTCATCACCACC 2948
QY 121 AGCACCCGCACCTGGGCGCTTGCCCACTTACATAACCACTCTACAAGCAAAATCTCCAGT 180
Db 2949 AGCACCCGCACCTGGGCGCTTGCCCACTTACATAACCACTCTACAAGCAAAATCTCCAGT 3008
QY 181 GCTTCAACGGGGGCCAGCAACGACAAACCACTACTTTCGGCTCAGACACCCCTCTGGGGGTAT 240
Db 3009 GCTTCAACGGGGGCCAGCAACGACAAACCACTACTTTCGGCTCAGACACCCCTCTGGGGGTAT 3068
QY 241 TTTGATTTCAACAGATTCCACCTGGCCTTTTCCACCGTACTGGCAGCGACTCATCAAC 300
Db 3069 TTTGATTTCAACAGATTCCACCTGGCCTTTTCCACCGTACTGGCAGCGACTCATCAAC 3128
QY 301 AACAAATTGGGGATTCCGGCCCAAGAGACTCAACTTTCAAACCTTTCAAACATCCAAGTCAAG 360
```

Db 3129 AACATTTGGGGAATTCGGGCCCAAGAGACTCAACTTCAAACCTCTTCAACATCCAAAGTCAAG 3188  
QY 361 GAGGTCAACGACGAATGATGGCGTCAACAACCATCGCTAAATAACCTTACCAGCAGCGTTCAA 420  
Db 3189 GAGGTCAACGACGAATGATGGCGTCAACAACCATCGCTAAATAACCTTACCAGCAGCGTTCAA 3248  
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTTCGATAGTCTCGGTCTCGGTCTCGGACACAGGCTGC 480  
Db 3249 GTCTTCTCGGACTCGGAGTACAGCTTTCGATAGTCTCGGTCTCGGTCTCGGACACAGGCTGC 3308  
QY 481 CTCCTCTCGGTCTCGGCGGAGCGTTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 540  
Db 3309 CTCCTCTCGGTCTCGGCGGAGCGTTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 3368  
QY 541 AATGGCAGCAACCCCGTGGACCTTTCATCTCTTTACTGCTGGAATATTTCCCTTCTCAG 600  
Db 3369 AATGGCAGCAACCCCGTGGACCTTTCATCTCTTTACTGCTGGAATATTTCCCTTCTCAG 3428  
QY 601 ATGCTGAGAACCGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660  
Db 3429 ATGCTGAGAACCGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 3488  
QY 661 AGCAGCTACGCGCACAGCAGCCTGGACCGGCTGATGATCCTCTCATCGACCAATAC 720  
Db 3489 AGCAGCTACGCGCACAGCAGCCTGGACCGGCTGATGATCCTCTCATCGACCAATAC 3548  
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTGCAACAAAGACTTGTCTG 780  
Db 3549 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTGCAACAAAGACTTGTCTG 3608  
QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAACTGGCTTACCTGACCC 840  
Db 3609 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAAACTGGCTTACCTGACCC 3668  
QY 841 TGTATTCCGACAGCGCGCTTCTTAAACAAACAAACAGACAAACAGCAATTTTACC 900  
Db 3669 TGTATTCCGACAGCGCGCTTCTTAAACAAACAAACAGACAAACAGCAATTTTACC 3728  
QY 901 TGGACTGGTGTCTTCAAAATATACTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 960  
Db 3729 TGGACTGGTGTCTTCAAAATATACTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 3788  
QY 961 GCTATGGCTTCAACAAAGACGACGAAAGTCTTTCCCATGAGCGGTGTCATGATT 1020  
Db 3789 GCTATGGCTTCAACAAAGACGACGAAAGTCTTTCCCATGAGCGGTGTCATGATT 3848  
QY 1021 TTTGGAAAAGAGAGCGCGGAGCTTCAACACATGCAATGGAACATGATGATTACAGAC 1080  
Db 3849 TTTGGAAAAGAGAGCGCGGAGCTTCAACACATGCAATGGAACATGATGATTACAGAC 3908  
QY 1081 GAAGAGGAATTAAGCCACTAACCTCTGGCCACCGAAGATTTGGACCGTGGCAGTC 1140  
Db 3909 GAAGAGGAATTAAGCCACTAACCTCTGGCCACCGAAGATTTGGACCGTGGCAGTC 3968  
QY 1141 AATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTGCATGCTATGGAGCATTA 1200  
Db 3969 AATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTGCATGCTATGGAGCATTA 4028  
QY 1201 CCTGGCATGTGTGGCAAGATAGAGACGTGTACTGTGAGGGTCCCATTTGGGCCAAAATT 1260  
Db 4029 CCTGGCATGTGTGGCAAGATAGAGACGTGTACTGTGAGGGTCCCATTTGGGCCAAAATT 4088  
QY 1261 CCTCACACATGACATTTTCAACCGCTCTCTTATGCGCGGCTTTGAGACTCAAGAAC 1320  
Db 4089 CCTCACACATGACATTTTCAACCGCTCTCTTATGCGCGGCTTTGAGACTCAAGAAC 4148  
QY 1321 CCGCTCTCAGATCTCTCATCAAAAACACGCTGTTCTCGAATCTCTCGGCGGAGTTT 1380  
Db 4149 CCGCTCTCAGATCTCTCATCAAAAACACGCTGTTCTCGAATCTCTCGGCGGAGTTT 4208  
QY 1381 TCAGCTACAAAGTTTGCTTCATTCACCCAACTCTCCACAGACAAGTGAAGTGGAA 1440  
Db 4209 TCAGCTACAAAGTTTGCTTCATTCACCCAACTCTCCACAGACAAGTGAAGTGGAA 4268

QY 1441 ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCCGCTGGAATCCCGAAGTGCAGTACACA 1500  
Db 4269 ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCCGCTGGAATCCCGAAGTGCAGTACACA 4328  
QY 1501 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATCGACTTTTATCT 1560  
Db 4329 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATCGACTTTTATCT 4388  
QY 1561 GAGCCTCGCCCCATTGGCACCCGTTTACCTTACCCGTCCTCTGTA 1604  
Db 4389 GAGCCTCGCCCCATTGGCACCCGTTTACCTTACCCGTCCTCTGTA 4432  
RESULT 9  
ADE76507  
ID ADE76507 standard; DNA; 4718 BP.  
XX  
AC ADE76507;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
Adeno-aassociated virus (AAV) related DNA, SEQ ID No 6.  
DE  
XX  
KW adeno-aassociated virus; AAV; cytostatic; antiposietic; antirheumatic;  
KW antiarthritic; neuroprotective; antidiabetic; antithyroid;  
KW dermatological; antiinflammatory; gene therapy; vaccine;  
KW hyperproliferative; cancer; psoriasis; autoimmune disease;  
KW rheumatoid arthritis; multiple sclerosis; diabetes;  
KW autoimmune thyroiditis; scleroderma; Crohn's disease; gene; ds.  
XX  
OS Adeno-aassociated virus 1.  
XX  
PN EP1310571-A2.  
XX  
PD 14-MAY-2003.  
XX  
PF 12-NOV-2002; 2002EP-00257826.  
XX  
PR 13-NOV-2001; 2001US-0350607P.  
PR 17-DEC-2001; 2001US-0341117P.  
PR 01-MAY-2002; 2002US-0377066P.  
PR 05-JUN-2002; 2002US-0386675P.  
XX  
PA (UYPE-) UNIV PENNSYLVANIA.  
XX  
PI Gao G, Wilson JM, Alvira M;  
XX  
XX WPI; 2003-450984/43.  
XX  
PT Detecting adeno-associated virus sequences in a sample, useful for e.g.  
PT preventing or treating hyperproliferative or autoimmune diseases,  
PT comprises subjecting a sample having a DNA to amplification via  
PT polymerase chain reaction.  
XX  
PS Claim 14; SEQ ID NO 6; 419pp; English.  
XX  
CC The invention relates to a novel method for detecting adeno-associated  
CC virus (AAV) sequences in a sample, which comprises subjecting a sample  
CC containing a DNA to amplification via a polymerase chain reaction (PCR).  
CC The AAV sequence have the following activities: cytostatic,  
CC antiposietic, antirheumatic, antiarthritic, neuroprotective,  
CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV  
CC sequence can be used in gene therapy or as part of a vaccine to treat  
CC disorders. The method is useful in detecting and/or identifying AAV  
CC sequences and isolating novel sequences that are identified. The  
CC sequences may be used e.g. for preventing or treating hyperproliferative  
CC conditions such as cancers and psoriasis, and other autoimmune diseases  
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune  
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence  
CC represents an AAV related DNA sequence of the invention.  
XX  
SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;



Query Match		100.0%;	Score 1604;	DB 10;	Length 4718;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1604;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGCTTACGGCGTGGCGACCAATGGCAGACAATAACGAAGGCGCGAGCGAGTGGGT	60		
Db	2829	ATGGCTTACGGCGTGGCGACCAATGGCAGACAATAACGAAGGCGCGAGCGAGTGGGT	2888		
Qy	61	AATGCTTCAGGAAATTTGGCAITTCGATTTCCATATGCTGGGGGACAGAGTCAATCAACC	120		
Db	2889	AATGCTTCAGGAAATTTGGCAITTCGATTTCCATATGCTGGGGGACAGAGTCAATCAACC	2948		
Qy	121	AGCACCGCACTGGGCTTGGCCACCTCAATTAACCACTCTCAAGCAAAATCTCCAGT	180		
Db	2949	AGCACCGCACTGGGCTTGGCCACCTCAATTAACCACTCTCAAGCAAAATCTCCAGT	3008		
Qy	181	GCTTCAACGGGGCCAGCAACACCACTACTTCGGCTACAGACCCCTGGGGGTAT	240		
Db	3009	GCTTCAACGGGGCCAGCAACACCACTACTTCGGCTACAGACCCCTGGGGGTAT	3068		
Qy	241	TTTGATTTCAACAGATTCACATGCCACTTTTCAACACGCTGACTGGCAGCACTCATCAAC	300		
Db	3069	TTTGATTTCAACAGATTCACATGCCACTTTTCAACACGCTGACTGGCAGCACTCATCAAC	3128		
Qy	301	AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG	360		
Db	3129	AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG	3188		
Qy	361	GAGGTCAACGAATGATGGCGTCAACACCATCGCTTAATACTTACAGACGGTTCAA	420		
Db	3189	GAGGTCAACGAATGATGGCGTCAACACCATCGCTTAATACTTACAGACGGTTCAA	3248		
Qy	421	GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGCTCTCGGCTCTCGCACCAAGGCTGC	480		
Db	3249	GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGCTCTCGGCTCTCGCACCAAGGCTGC	3308		
Qy	481	CTCCCTCGCTTCGGCGGACGCTGTTTCATGATTCGCAATACGCTCACTGAGCTCAAC	540		
Db	3309	CTCCCTCGCTTCGGCGGACGCTGTTTCATGATTCGCAATACGCTCACTGAGCTCAAC	3368		
Qy	541	AATGGCAGCAAGCGTGGGACGTTTCATCTTTTACTGCTGGAATATTTCTCTCTCAG	600		
Db	3369	AATGGCAGCAAGCGTGGGACGTTTCATCTTTTACTGCTGGAATATTTCTCTCTCAG	3428		
Qy	601	ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC	660		
Db	3429	ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC	3488		
Qy	661	AGCAGCTACGGGACAGCCAGAGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC	720		
Db	3489	AGCAGCTACGGGACAGCCAGAGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC	3548		
Qy	721	CTGTATTACTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAGAGACTTGCTG	780		
Db	3549	CTGTATTACTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAGAGACTTGCTG	3608		
Qy	781	TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAAACTGGCTACTCGACCC	840		
Db	3609	TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAAACTGGCTACTCGACCC	3668		
Qy	841	TGTTATCGGACGAGCGCGTTTCAAAACAAAAAGACAAACAGACGAATTTTACC	900		
Db	3669	TGTTATCGGACGAGCGCGTTTCAAAACAAAAAGACAAACAGACGAATTTTACC	3728		
Qy	901	TGGACTGGGTCTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT	960		
Db	3729	TGGACTGGGTCTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT	3788		
Qy	961	GCTATGGCTCACAAAAAGACGACGAGA CAAGTCTTTTCCCATGACGGGTGTGATTT	1020		
Db	3789	GCTATGGCTCACAAAAAGACGACGAGA CAAGTCTTTTCCCATGACGGGTGTGATTT	3848		

Qy	1021	TTTGGAAAGAGAGCGCGGAGCTTCAAAACACTGCATTTGGCAATGTGCATGATTACAGAC	1080		
Db	3849	TTTGGAAAGAGAGCGCGGAGCTTCAAAACACTGCATTTGGCAATGTGCATGATTACAGAC	3908		
Qy	1081	GAAGAGGAAATTAAGAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC	1140		
Db	3909	GAAGAGGAAATTAAGAGCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC	3968		
Qy	1141	AATTTCCAGACAGCAGACACAGACCTTCGACCGGAGATGTGCATTTGGGAGCATTA	1200		
Db	3969	AATTTCCAGACAGCAGACACAGACCTTCGACCGGAGATGTGCATTTGGGAGCATTA	4028		
Qy	1201	CCTGGCATGGTGTGGCAAGATAGAGCGTGTACCTGACGGGTCCCATTTGGGCCCAAAAT	1260		
Db	4029	CCTGGCATGGTGTGGCAAGATAGAGCGTGTACCTGACGGGTCCCATTTGGGCCCAAAAT	4088		
Qy	1261	CCTCACACAGATGACACACTTTTCAACCGCTCTCTTATGGCGGGCTTTGGGACTCAAGAAC	1320		
Db	4089	CCTCACACAGATGACACACTTTTCAACCGCTCTCTTATGGCGGGCTTTGGGACTCAAGAAC	4148		
Qy	1321	CGGCTCTCAGATCTCTCATCAAAACACGCTGTTCCTGCGAAATCTCTCCGGCGAGTTT	1380		
Db	4149	CGGCTCTCAGATCTCTCATCAAAACACGCTGTTCCTGCGAAATCTCTCCGGCGAGTTT	4208		
Qy	1381	TCAGCTACAAAGTTTGTCTTCAATTCATCACCACTATCCAGGACCAAGTGTGTGAA	1440		
Db	4209	TCAGCTACAAAGTTTGTCTTCAATTCATCACCACTATCCAGGACCAAGTGTGTGAA	4268		
Qy	1441	ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGAAAGTGCAGTACACA	1500		
Db	4269	ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGAAAGTGCAGTACACA	4328		
Qy	1501	TCCAATTAACAAATCTGCCAACGTTGATTTTACTGTGGAACAATAATGGACTTTATACT	1560		
Db	4329	TCCAATTAACAAATCTGCCAACGTTGATTTTACTGTGGAACAATAATGGACTTTATACT	4388		
Qy	1561	GAGCTTCGCCCATTTGGCACCGGTTACCTTACCCGTCCCTGTGA	1604		
Db	4389	GAGCTTCGCCCATTTGGCACCGGTTACCTTACCCGTCCCTGTGA	4432		
RESULT 10					
ADL13984					
ID	ADL13984 standard; DNA; 4718 BP.				
XX	ADL13984;				
AC	AC				
DT	DT				
XX	06-MAY-2004 (first entry)				
DE	Adeno-associated virus serotype 1 complete DNA.				
XX	ss; cytostatic; neuroprotective; antiinflammatory; gene therapy;				
KW	expression construct; adeno-associated virus;				
KW	integration efficiency element; inverted terminal repeat; integration;				
KW	chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;				
KW	retinoblastoma; inflammatory disease; arthritis;				
KW	neurodegenerative disease.				
OS	Adeno-associated virus 1.				
XX	WO2003087334-A2.				
PN	23-OCT-2003.				
PD	09-APR-2003; 2003WO-US011191.				
PF	09-APR-2002; 2002US-0371044P.				
PR	(CORR ) CORNELL RES FOUND INC.				
PA	Falck-Pedersen ES, Philpott N;				
XX	WPI; 2003-833723/77.				



XX New expression construct comprising a nucleic acid sequence encoding an  
PT adeno-associated virus integration efficiency element, useful for  
PT treating cancer e.g. lung cancer or colon cancer or inflammatory disease  
PT e.g. arthritis.  
XX

PS Disclosure; SEQ ID NO 3; 62pp; English.

XX  
CC The invention relates to an expression construct comprising a nucleic  
CC acid sequence encoding an adeno-associated virus integration efficiency  
CC element (AAV IRE), which is devoid of AAV inverted terminal repeats (AAV  
CC ITRs) and site-specifically integrates into a host cell chromosome when  
CC provided to the host cell in conjunction with an AAV Rep protein. The  
CC expression construct can be used as a therapeutic factor for treating a  
CC mammal for a pathologic state which is cancer, including lung cancer,  
CC colon cancer, renal cancer, anal cancer, bile duct cancer, bladder  
CC cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,  
CC cervical cancer, lymphoma, endometrial cancer, esophageal cancer,  
CC gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,  
CC liver cancer, multiple myeloma, neuroblastoma, ovarian cancer, pancreatic  
CC cancer, prostatic cancer, retinoblastoma, skin cancer, stomach cancer,  
CC testicular cancer, thymus cancer or thyroid cancer. Other pathologic  
CC state includes inflammatory disease (arthritis), neurodegenerative  
CC disease, a disease of an organ attributed to the presence of increased or  
CC decreased level of a particular gene product(s). This sequence  
CC corresponds to the AAV serotype 1 complete DNA including the IRE  
CC sequence.

SQ Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;

Query Match 100.0%; Score 1604; DB 10; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCAGGCGTGGCGACCAATGGCAGACAATAACGAGCGCCGACGAGTGGGT 60  
DB 2829 ATGGCTTCAGGCGTGGCGACCAATGGCAGACAATAACGAGCGCCGACGAGTGGGT 2888  
QY 61 AATGCTTCAGGAAATGGCAATGCGATTCCACATGCTGGCGACAGAGTCATCACACC 120  
DB 2889 AATGCTTCAGGAAATGGCAATGCGATTCCACATGCTGGCGACAGAGTCATCACACC 2948  
QY 121 AGCACCCGACCTGGCGCTTGGCCACCTAATAAACCACCTCTACAAAGCAAAATCTCCAGT 180  
DB 2949 AGCACCCGACCTGGCGCTTGGCCACCTAATAAACCACCTCTACAAAGCAAAATCTCCAGT 3008  
QY 181 GCTTCAACGGGGCCAGCAACGACCAACCTACTTCCGGCTACAGACCCCTGGGGGTAT 240  
DB 3009 GCTTCAACGGGGCCAGCAACGACCAACCTACTTCCGGCTACAGACCCCTGGGGGTAT 3068  
QY 241 TTTGATTTCAACAGATTCCCACTGCCACTTTTACCACGCTGACTGGCAGGACTCATCAAC 300  
DB 3069 TTTGATTTCAACAGATTCCCACTGCCACTTTTACCACGCTGACTGGCAGGACTCATCAAC 3128  
QY 301 AACAAATGGGGATTCGGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360  
DB 3129 AACAAATGGGGATTCGGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 3188  
QY 361 GAGGTCAACGAGAAATGATGGCGTCAACCAATCGCTAATAAACCCTTACGACGCGTTCAA 420  
DB 3189 GAGGTCAACGAGAAATGATGGCGTCAACCAATCGCTAATAAACCCTTACGACGCGTTCAA 3248  
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCGATGATTTCCGCAATACGGCTACCTGACGCTCAAC 480  
DB 3249 GTCTTCTCGGACTCGGAGTACAGCTTCGATGATTTCCGCAATACGGCTACCTGACGCTCAAC 3308  
QY 481 CTCCTCTCCGTTCCCGGCGGAGTGTTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 540  
DB 3309 CTCCTCTCCGTTCCCGGCGGAGTGTTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 3368  
QY 541 AATGGCAGCAACCGGTGGAGCTTCACTCTTTTACTGCTGGAATATTTCCCTTCTCAG 600  
DB 3369 AATGGCAGCAACCGGTGGAGCTTCACTCTTTTACTGCTGGAATATTTCCCTTCTCAG 3428

RESULT 11  
ADG39758  
ID ADG39758 standard; DNA; 4718 BP.

QY 601 ATGCTGAGAACGGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660  
DB 3429 ATGCTGAGAACGGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 3488  
QY 661 AGCAGCTACGGCGCACAGCCAGAGCGCTGGACCGGCTGATGAATCCCTCTCATCGCAACATAC 720  
DB 3489 AGCAGCTACGGCGCACAGCCAGAGCGCTGGACCGGCTGATGAATCCCTCTCATCGCAACATAC 3548  
QY 721 CTGTATTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCACCAAAACAGGACTTGTCTG 780  
DB 3549 CTGTATTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCACCAAAACAGGACTTGTCTG 3608  
QY 781 TTTAGCCGTGGTCTTCAGCTGGCATGCTGTTCAGCCCAAAACCTGGCTACTTGGACCC 840  
DB 3609 TTTAGCCGTGGTCTTCAGCTGGCATGCTGTTCAGCCCAAAACCTGGCTACTTGGACCC 3668  
QY 841 TGTATTCCGCGACGCGGCTTTCTAAACAAAAACAGACCAACAAACAGCAATTTTACC 900  
DB 3669 TGTATTCCGCGACGCGGCTTTCTAAACAAAAACAGACCAACAAACAGCAATTTTACC 3728  
QY 901 TGGACTGGTGTCTTCAAAATATAAATCAATGGGCGTGAATCCATCATCAACCTGGCACT 960  
DB 3729 TGGACTGGTGTCTTCAAAATATAAATCAATGGGCGTGAATCCATCATCAACCTGGCACT 3788  
QY 961 GCTATGGCTCTACACAAAGACGACGAAGTGTCTTCCATGAGCGGTGTCTGATTT 1020  
DB 3789 GCTATGGCTCTACACAAAGACGACGAAGTGTCTTCCATGAGCGGTGTCTGATTT 3848  
QY 1021 TTTGAAAAGAGAGCGCGGAGCTTCAAACTGCTGATTTGGCAATGTCATGATTAACAGAC 1080  
DB 3849 TTTGAAAAGAGAGCGCGGAGCTTCAAACTGCTGATTTGGCAATGTCATGATTAACAGAC 3908  
QY 1081 GAAGAGGAAATTAAGAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140  
DB 3909 GAAGAGGAAATTAAGAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 3968  
QY 1141 AATTTCCAGAGCAGCAGCAGACAGCCCTGCGACCGAGAGTGTGCTATGAGGAGCATTTA 1200  
DB 3969 AATTTCCAGAGCAGCAGCAGCAGACAGCCCTGCGACCGAGAGTGTGCTATGAGGAGCATTTA 4028  
QY 1201 CCTGCGATGGTGTGGCAAGATAGACGCTGATGCTGAGGCTCCCAATTTGGGCCAAAATTT 1260  
DB 4029 CCTGCGATGGTGTGGCAAGATAGACGCTGATGCTGAGGCTCCCAATTTGGGCCAAAATTT 4088  
QY 1261 CCTCAACAGATGGGACACTTTTCAACCGCTCTCTCTTATGGGCGGCTTTTGGACTCAAGAAC 1320  
DB 4089 CCTCAACAGATGGGACACTTTTCAACCGCTCTCTCTTATGGGCGGCTTTTGGACTCAAGAAC 4148  
QY 1321 CGCGCTCTCTCAGATCCTCATCAAAACACGCGCTGTTCCTGCGAATCTCTCCGCGGAGTTT 1380  
DB 4149 CGCGCTCTCTCAGATCCTCATCAAAACACGCGCTGTTCCTGCGAATCTCTCCGCGGAGTTT 4208  
QY 1381 TCAGCTAACAAGTTTGTCTTCATTCACCCAAATCTCCACAGGACCAAGTGTGAGTGGAA 1440  
DB 4209 TCAGCTAACAAGTTTGTCTTCATTCACCCAAATCTCCACAGGACCAAGTGTGAGTGGAA 4268  
QY 1441 ATTGAATGGGAGCTCGAGAAACAAACAGCAACGCTGGAATCCCGAAGTGCAGTACACA 1500  
DB 4269 ATTGAATGGGAGCTCGAGAAACAAACAGCAACGCTGGAATCCCGAAGTGCAGTACACA 4328  
QY 1501 TCCAAATATGCAAAATCTGCCAAACGTTGATTTTACTGTGGCAACCAATGGACTTTTACT 1560  
DB 4329 TCCAAATATGCAAAATCTGCCAAACGTTGATTTTACTGTGGCAACCAATGGACTTTTACT 4388  
QY 1561 GAGCCTCGCCCCATTTGGCAGCCGTTACTTACCCGTCCTCTGTA 1604  
DB 4389 GAGCCTCGCCCCATTTGGCAGCCGTTACTTACCCGTCCTCTGTA 4432





Db 6127 TGGAGTGGTCTTCAAAATATTAACCTCAATGGCGTGAATCATCATCAACCTGGCACT 6186  
Qy 961 GCTATGGCTCTACACAAAGACGACGAAGCAAGTCTTTCCCATGAGCGGTGTCATGATT 1020  
Db 6187 GCTATGGCTCTACACAAAGACGACGAAGCAAGTCTTTCCCATGAGCGGTGTCATGATT 6246  
Qy 1021 TTTGGAAAAGAGCGCCGGAGCTTCAAAACATGCAATTTGGACAAATGTCATGATTACAGAC 1080  
Db 6247 TTTGGAAAAGAGCGCCGGAGCTTCAAAACATGCAATTTGGACAAATGTCATGATTACAGAC 6306  
Qy 1081 GAAGAGGAAATTAAGCCACTTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140  
Db 6307 GAAGAGGAAATTAAGCCACTTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 6366  
Qy 1141 AATTTCCAGACGACGACGACAGACCTCGGACCGGAGATGTGCATGCTATGGGAGCATTA 1200  
Db 6367 AATTTCCAGACGACGACGACAGACCTCGGACCGGAGATGTGCATGCTATGGGAGCATTA 6426  
Qy 1201 CTTGGCATGTGTGGCAAGATAGAGACGTTGACCTGACGGGTCCCATTTGGGCGCAAAATT 1260  
Db 6427 CTTGGCATGTGTGGCAAGATAGAGACGTTGACCTGACGGGTCCCATTTGGGCGCAAAATT 6486  
Qy 1261 CTTACACAGATGGACACTTTTACCCGCTCTCTCTTTATGGGGGCTTTGGACTCAAGAAC 1320  
Db 6487 CTTACACAGATGGACACTTTTACCCGCTCTCTCTTTATGGGGGCTTTGGACTCAAGAAC 6546  
Qy 1321 CGCCCTCTCAGATCCTCATCAAAAACAGCCCTGTCTCTCGGAATCCTCCGGGGAGTTT 1380  
Db 6547 CGCCCTCTCAGATCCTCATCAAAAACAGCCCTGTCTCTCGGAATCCTCCGGGGAGTTT 6606  
Qy 1381 TCAGCTACAAAGTTGCTTCAATTCATCAACCAATCTCCACAGGACAAAGTGTGGAA 1440  
Db 6607 TCAGCTACAAAGTTGCTTCAATTCATCAACCAATCTCCACAGGACAAAGTGTGGAA 6666  
Qy 1441 ATTGAATGGAGCTGCAGAAAGAAAACAGACGCTGGAAATCCCGAAGTGCAGTACACA 1500  
Db 6667 ATTGAATGGAGCTGCAGAAAGAAAACAGACGCTGGAAATCCCGAAGTGCAGTACACA 6726  
Qy 1501 TCCAAATATGCAAAATCTGCCAAGTGTGATTTTACTGTGGACAAATGGACTTTATCT 1560  
Db 6727 TCCAAATATGCAAAATCTGCCAAGTGTGATTTTACTGTGGACAAATGGACTTTATCT 6786  
Qy 1561 GAGCCTGGCCCAATGGCACCCGTTACTTACCCGTCCTCTGTA 1604  
Db 6787 GAGCCTGGCCCAATGGCACCCGTTACTTACCCGTCCTCTGTA 6830

RESULT 13

ADZ26931 standard; DNA; 2211 BP.

ID ADZ26931

XX AC ADZ26931;

XX AC ADZ26931;

DT 30-JUN-2005 (first entry)

XX DE Adeno-associated virus DNA SEQ ID NO 81.

XX KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;  
XX KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;  
XX KW bacterial infection; cancer; ulcerative colitis; antirheumatic;  
XX KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;  
XX KW antipeptidic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;  
XX KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.

XX OS Adeno-associated virus.

XX OS Adeno-associated virus.

XX PN W0205033321-A2.

XX PD 14-APR-2005.

XX PF 30-SEP-2004; 2004MO-US028817.

XX XX

PR 30-SEP-2003; 2003US-0508226P.  
PR 29-APR-2004; 2004US-0566546P.  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX PI Wilson JM, Gao G, Alvira MR, Vandenbergh LH;  
XX WPI; 2005-285437/29.  
XX New adeno-associated virus (AAV) clade comprising at least three AAV  
PT members, useful for preventing and/or treating arthritis, multiple  
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial  
PT infection and cancer.  
XX Claim 19; SEQ ID NO 81; 569pp; English.  
XX The invention relates to an adeno-associated virus (AAV) clade comprising  
CC at least three AAV members, where each member of the AAV clade is  
CC phylogenetically related as determined using a Neighbor-Joining heuristic  
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson  
CC correction distance measurement of no more than 0.05. The methods and  
CC compositions of the present invention are useful for the prevention  
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,  
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's  
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative  
CC colitis. The present sequence represents an adeno-associated virus DNA.  
XX SQ Sequence 2211 BP; 582 A; 653 C; 535 G; 441 T; 0 U; 0 Other;

Query Match 99.8%; Score 1600.8; DB 14; Length 2211;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1602; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAAATAACGAAGGCGCGAGTGGGT 60  
Db 607 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAAATAACGAAGGCGCGAGTGGGT 666  
Qy 61 AATGCTCAGGAATTTGGCATTTGGATTCACATGGCTGGGGGACAGAGTATCAACACC 120  
Db 667 AATGCTCAGGAATTTGGCATTTGGATTCACATGGCTGGGGGACAGAGTATCAACACC 726  
Qy 121 AGCACCCGACCTGGGCGCTTGGCCACCTACAAATACCACTCTACAAGCAAAATCTCCAGT 180  
Db 727 AGCACCCGACCTGGGCGCTTGGCCACCTACAAATACCACTCTACAAGCAAAATCTCCAGT 786  
Qy 181 GCTTCAACGGGGCGCAGCAACCACTACTTTGGGCTACAGCACCCCTGGGGGGTAT 240  
Db 787 GCTTCAACGGGGCGCAGCAACCACTACTTTGGGCTACAGCACCCCTGGGGGGTAT 846  
Qy 241 TTTGATTTCAACAGATTCACATGGCCACTTTTCAACACATGGTGGCAGGACTCATCAAC 300  
Db 847 TTTGATTTCAACAGATTCACATGGCCACTTTTCAACACATGGTGGCAGGACTCATCAAC 906  
Qy 301 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTTCAAACCTTTCAAACATCCAAGTCAAG 360  
Db 907 AACAAATGGGGATTCGGGCCCAAGAGACTCAACTTTCAAACCTTTCAAACATCCAAGTCAAG 966  
Qy 361 GAGGTCAACGACGAATGATGGCGTCAACACCACTCGCTAATAACCTTTACCAGCACGGTTCAA 420  
Db 967 GAGGTCAACGACGAATGATGGCGTCAACACCACTCGCTAATAACCTTTACCAGCACGGTTCAA 1026  
Qy 421 GTCTTCTGGAGTCCGAGTACAGATTTCGTTACGTTCTCGGCTCTGGCGACACGAGGCTGC 480  
Db 1027 GTCTTCTGGAGTCCGAGTACAGATTTCGTTACGTTCTCGGCTCTGGCGACACGAGGCTGC 1086  
Qy 481 CTCCCTCGTTCCCGCGGACGTTTCATGTTCCGCAATACGGCTACTGACGCTCAAC 540  
Db 1087 CTCCCTCGTTCCCGCGGACGTTTCATGTTTCGCAATACGGCTACTGACGCTCAAC 1146  
Qy 541 AATGGCAGCAAGCGGTGGAGCGTTTCATCTTTTACTGCTTGAATAATTTCCCTTTCTCAG 600  
Db 1147 AATGGCAGCAAGCGGTGGAGCGTTTCATCTTTTACTGCTTGAATAATTTCCCTTTCTCAG 1206

QY 601 ATGCTGAGAAAGCGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCTTTCCAC 660  
DB 1207 ATGCTGAGAAAGCGGCAACAACTTTACCTTCAGCTACACCTTTGAGGAAGTGCTTTCCAC 1266  
QY 661 AGCAGCTAGCGGACAGCGAGCTGGACCGGCTGATGAATCCTTCATCGACCAATAC 720  
DB 1267 AGCAGCTAGCGGACAGCGAGCTGGACCGGCTGATGAATCCTTCATCGACCAATAC 1326  
QY 721 CTGTATTACTGAACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAACAGGACTTTGCTG 780  
DB 1327 CTGTATTACCGNACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAACAGGACTTTGCTG 1386  
QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAATCTGGCTACCTGACCC 840  
DB 1387 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAATCTGGCTACCTGACCC 1446  
QY 841 TGTATTGGGACAGCGCGTCTTAAACAAACAAACAGACACAAACAGCAATTTTACC 900  
DB 1447 TGTATTGGGACAGCGCGTCTTAAACAAACAAACAGACACAAACAGCAATTTTACC 1506  
QY 901 TGGACTGGTGTCTCAAAATATATACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 960  
DB 1507 TGGACTGGTGTCTCAAAATATATACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 1566  
QY 961 GCTATGGCCTCACACAAAGACGACGAAGCAAGTCTTTCCCATGAGCGGTGTCATGATT 1020  
DB 1567 GCTATGGCCTCACACAAAGACGACGAAGCAAGTCTTTCCCATGAGCGGTGTCATGATT 1626  
QY 1021 TTTGAAAAAGAGCGCGGAGCTTCAACACTGCAATTTGGACAAATGTCATGATTACAGAC 1080  
DB 1627 TTTGAAAAAGAGCGCGGAGCTTCAACACTGCAATTTGGACAAATGTCATGATTACAGAC 1686  
QY 1081 GAAGAGGAATTAAGCCCTAATACCTGTGGCCACCGAAGATTTCGGACCGTGGCAGTC 1140  
DB 1687 GAAGAGGAATTAAGCCCTAATACCTGTGGCCACCGAAGATTTCGGACCGTGGCAGTC 1746  
QY 1141 AATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTGCTATGCGGAGCATTA 1200  
DB 1747 AATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTGCTATGCGGAGCATTA 1806  
QY 1201 CTGGCATGGTGTGGCAAGATAGAGACGTTGACCTGACAGGTCCTATTTGGGCCAAATTT 1260  
DB 1807 CTGGCATGGTGTGGCAAGATAGAGACGTTGACCTGACAGGTCCTATTTGGGCCAAATTT 1866  
QY 1261 CCTCACAGATGGACATTTACCCGTCTCCTTTATGCGCGGCTTTGGACCTCAAGAAC 1320  
DB 1867 CCTCACAGATGGACATTTACCCGTCTCCTTTATGCGCGGCTTTGGACCTCAAGAAC 1926  
QY 1321 CCGCCTCCTCAGATCCTCATCAAAACACACGCTGTTCTCGGAATCCTCGGCGGAGTTT 1380  
DB 1927 CCGCCTCCTCAGATCCTCATCAAAACACACGCTGTTCTCGGAATCCTCGGCGGAGTTT 1986  
QY 1381 TCAGCTACAAAGTTTGCTTCATTCATCCCAATATCTCCAAGGACAAGTGAGTGGAA 1440  
DB 1987 TCAGCTACAAAGTTTGCTTCATTCATCCCAATATCTCCAAGGACAAGTGAGTGGAA 2046  
QY 1441 ATTGAATGGAGCTGCAGAAAGAAACAGACGCGTGGAAATCCGAGTGGCAGTACACA 1500  
DB 2047 ATTGAATGGAGCTGCAGAAAGAAACAGACGCGTGGAAATCCGAGTGGCAGTACACA 2106  
QY 1501 TCCAAATTATCAAAATCTGCAACGTTGATTTTACTGTGGACAAACATGACCTTTTACT 1560  
DB 2107 TCCAAATTATCAAAATCTGCAACGTTGATTTTACTGTGGACAAACATGACCTTTTACT 2166  
QY 1561 GAGCCTCGCCCCATTTGGCACCCGTTACCTTTACCCGTCCTCTGTA 1604  
DB 2167 GAGCCTCGCCCCATTTGGCACCCGTTACCTTTACCCGTCCTCTGTA 2210

RESULT 14  
ADZ26930  
ID ADZ26930 standard; DNA; 2214 BP.  
XX

AC ADZ26930;  
XX 30-JUN-2005 (first entry)  
XX Adeno-associated virus DNA SEQ ID NO 80.  
DE rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;  
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;  
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;  
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;  
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;  
KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.  
XX Adeno-associated virus.  
OS WO2005033321-A2.  
XX 14-APR-2005.  
XX 30-SEP-2004; 2004WO-US028817.  
XX 30-SEP-2003; 2003US-0508226P.  
XX 29-APR-2004; 2004US-0566546P.  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;  
XX WPI; 2005-285437/29.  
XX New adeno-associated virus (AAV) clade comprising at least three AAV  
PT members, useful for preventing and/or treating arthritis, multiple  
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial  
PT infection and cancer.  
XX Claim 50; SEQ ID NO 80; 569pp; English.  
XX The invention relates to an adeno-associated virus (AAV) clade comprising  
CC at least three AAV members, where each member of the AAV clade is  
CC phylogenetically related as determined using a Neighbor-Joining heuristic  
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson  
CC correction distance measurement of no more than 0.05. The methods and  
CC compositions of the present invention are useful for the prevention  
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,  
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's  
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative  
CC colitis. The present sequence represents an adeno-associated virus DNA.  
XX  
SQ Sequence 2214 BP; 572 A; 658 C; 543 G; 441 T; 0 U; 0 Other;  
Query Match 99.4%; Score 1594.4; DB 14; Length 2214;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1598; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 ATGGCTTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCCGACGAGTGGGT 60  
DB 610 ATGGCTTCAGGCGGTGGCGCTCCCAATGGCAGACAATAACGAAGCGCGGAGTGGGT 669  
QY 61 AATGCTTCAGGAATTTGGGATTTGGGATTTCCATGCTGGGCGCAGAGTCAACACC 120  
DB 670 AATGCTTCAGGAATTTGGGATTTGGGATTTCCATGCTGGGCGCAGAGTCAACACC 729  
QY 121 AGCACCGGACCTGGGCGCTTGGCCACCTCAATTAACACCTCTACAGCAAAATCTCCAGT 180  
DB 730 AGCACCGGACCTGGGCGCTTGGCCACCTCAATTAACACCTCTACAGCAAAATCTCCAGT 789  
QY 181 GCTTCAACCGGGGCGCAGCAACCACTACTTCTCGGCTACAGCAACCCCTCGGGGTAT 240  
DB 790 GCTTCAACCGGGGCGCAGCAACCACTACTTCTCGGCTACAGCAACCCCTCGGGGTAT 849  
QY 241 TTTGATTCAACAGATTCCACTGCCACTTTTCAACAGTGACTGGGAGGACTCATCAAC 300  
DB 850 TTTGATTCAACAGATTCCACTGCCACTTTTCAACAGTGACTGGGAGGACTCATCAAC 909



QY 301 AACAAATGGGGATTCGGGCCAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360  
Db 910 AACAAATGGGGATTCGGGCCAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 969  
QY 361 GAGGTCAACAGCAATGATGGCGTCACAAACATCGCTAAATAACCTTACACGACCGGTTCAA 420  
Db 970 GAGGTCAACAGCAATGATGGCGTCACAAACATCGCTAAATAACCTTACACGACCGGTTCAA 1029  
QY 421 GTCTTCTCGGACTCGGAGTACAGAGTTCCGAGTTCCTCGGCTCTGGCTTCGGGACCAAGGGCTGC 480  
Db 1030 GTCTTCTCGGACTCGGAGTACAGAGTTCCGAGTTCCTCGGCTCTGGGACCAAGGGCTGC 1089  
QY 481 CTCCCTCCGTTCCCGCGGAGCGTGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 540  
Db 1090 CTCCCTCCGTTCCCGCGGAGCGTGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 1149  
QY 541 AATGGCAGCAAGCCCGTGGAGCGTTTCATCTCTTTACTGCTCGAATATTTCCCTTCTCAG 600  
Db 1150 AATGGCAGCAAGCCCGTGGAGCGTTTCATCTCTTTACTGCTCGAATATTTCCCTTCTCAG 1209  
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660  
Db 1210 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 1269  
QY 661 AGCAGCTACGCGACAGCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720  
Db 1270 AGCAGCTACGCGACAGCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGTCCAAATAC 1329  
QY 721 CTGTATTACTCTGAACAGAACTCAAAATCAGTCGGGAAGTGGCCAAACAAAGGACTTGTCTG 780  
Db 1330 CTGTATTACTCTGAACAGAACTCAAAATCAGTCGGGAAGTGGCCAAACAAAGGACTTGTCTG 1389  
QY 781 TTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTCAGGCCCAAAATCTGGCTACTCGACCC 840  
Db 1390 TTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTCAGGCCCAAAATCTGGCTACTCGACCC 1449  
QY 841 TGTATTCCGACAGCGCGTCTTCAAAACAAACAGACAAACAAACAGCAATTTTACC 900  
Db 1450 TGTATTCCGACAGCGCGTCTTCAAAACAAACAGACAAACAAACAGCAATTTTACC 1509  
QY 901 TGGACTGGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCAATCAACCTGGCACT 960  
Db 1510 TGGACTGGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCAATCAACCTGGCACT 1569  
QY 961 GCTATGGCTCTACACAAAGACGACGAAGCAAGTTCTTTCCCATGAGCGGTGTCATGATT 1020  
Db 1570 GCTATGGCTCTACACAAAGACGACGAAGCAAGTTCTTTCCCATGAGCGGTGTCATGATT 1629  
QY 1021 TTTGGAAAGAGAGCGCGGAGCTTCAAACTGCTATGGACAAATGTCATGATTCAGAC 1080  
Db 1630 TTTGGAAAGAGAGCGCGGAGCTTCAAACTGCTATGGACAAATGTCATGATTCAGAC 1689  
QY 1081 GAAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140  
Db 1690 GAAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1749  
QY 1141 AATTTCCAGACGACGACAGACCTTCGACCGGAGATGTGCTATGAGGACATTA 1200  
Db 1750 AATTTCCAGACGACGACAGACCTTCGACCGGAGATGTGCTATGAGGACATTA 1809  
QY 1201 CCTGGCATGGTGGCAAGATAGAGAGTGTACTCGAGGTTCCCATTTGGGCCAAATTT 1260  
Db 1810 CCTGGCATGGTGGCAAGATAGAGAGTGTACTCGAGGTTCCCATTTGGGCCAAATTT 1869  
QY 1261 CCTCACACAGATGACACTTTTACCCGCTCTCTCTTATGGCGGCTTTGAGCTCAAGAAC 1320  
Db 1870 CCTCACACAGATGACACTTTTACCCGCTCTCTCTTATGGCGGCTTTGAGCTCAAGAAC 1929  
QY 1321 CCGCCTCTCAGATCCTCATCAAAACACAGCCTGTTTCTGCGAATCCTCGGCGGAGTTT 1380  
Db 1930 CCGCCTCTCAGATCCTCATCAAAACACAGCCTGTTTCTGCGAATCCTCGGCGGAGTTT 1989

QY 1381 TCAGCTACAAAGTTTGTCTTCATCCACCAATACTCCACAGGACAAAGTGTGTGGAA 1440  
Db 1990 TCAGCTACAAAGTTTGTCTTCATCCACCAATACTCCACAGGACAAAGTGTGTGGAA 2049  
QY 1441 ATTGAATGGGAGCTGCAGAGAAAGAAACAGCAAGCGCTGGAATCCCGAAAGTGCAGTACACA 1500  
Db 2050 ATTGAATGGGAGCTGCAGAGAAAGAAACAGCAAGCGCTGGAATCCCGAAAGTGCAGTACACA 2109  
QY 1501 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTATCT 1560  
Db 2110 TCCAAATTATGCAAAATCTGCCAGCGTTGATTTTACTGTGGACAAACAATGGACTTTATCT 2169  
QY 1561 GAGCCTCGCCCATTTGGCACCGTTTACTTACCCTGCCCTGTA 1604  
Db 2170 GAGCCTCGCCCATTTGGCACCGTTTACTTACCCTGCCCTGTA 2213  
RESULT 15  
ADZ26928  
ID ADZ26928 standard; DNA; 2211 BP.  
XX  
AC ADZ26928;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE Adeno-associated virus DNA SEQ ID NO 78.  
XX  
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;  
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;  
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;  
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;  
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;  
KW virucide; antibacterial; cytostatic; antitumor; dermatological; ds.  
XX  
OS Adeno-associated virus.  
XX  
PN WO2005033321-A2.  
XX  
PD 14-APR-2005.  
XX  
PF 30-SEP-2004; 2004WO-US028817.  
XX  
PR 30-SEP-2003; 2003US-0508226P.  
PR 29-APR-2004; 2004US-0566546P.  
XX  
PA (UYPE-) UNIV PENNSYLVANIA.  
XX  
PI Wilson JM, Gao G, Alvira MR, Vandenberghe LH;  
XX  
PS WPI; 2005-285437/29.  
XX  
PT New adeno-associated virus (AAV) clade comprising at least three AAV  
members, useful for preventing and/or treating arthritis, multiple  
sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial  
infection and cancer.  
XX  
PS Claim 19; SEQ ID NO 78; 569pp; English.  
XX  
CC The invention relates to an adeno-associated virus (AAV) clade comprising  
at least three AAV members, where each member of the AAV clade is  
CC phylogenetically related as determined using a Neighbor-Joining heuristic  
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson  
CC correction distance measurement of no more than 0.05. The methods and  
CC compositions of the present invention are useful for the prevention  
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,  
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's  
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative  
CC colitis. The present sequence represents an adeno-associated virus DNA.  
XX  
SQ Sequence 2211 BP; 575 A; 654 C; 541 G; 441 T; 0 U; 0 Other;  
Query Match 99.3%; Score 1592.8; DB 14; Length 2211;  
Best Local Similarity 99.6%; Pred. No. 0;

Matches 1597; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ATGGCTTCAGGCGGTGGCCACCAATATGGCAGACAATAACGAAGGCGCGCAGGAGTGGGT 60  
DB |||||  
QY 607 ATGGCTTCAGGCGGTGGCCACCAATATGGCAGACAATAACGAAGGCGCGCAGGAGTGGGT 666  
DB |||||  
QY 61 AATGCTCAGGAAATTTGGCATTTCCGATTTCCCATGCTGGGCGCAGAGTGCATCAACC 120  
DB |||||  
QY 667 AATGCTCAGGAAATTTGGCATTTCCGATTTCCCATGCTGGGCGCAGAGTGCATCAACC 726  
DB |||||  
QY 121 AGCACCGCAGCTGGGCGCTGGCCACCTACATTAACCACTCTACAGCAAAATCTCCAGT 180  
DB |||||  
QY 727 AGCACCGCAGCTGGGCGCTGGCCACCTACATTAACCACTCTACAGCAAAATCTCCAGT 786  
DB |||||  
QY 181 GCTTCAACGGGCGCCAGCAACAGCAACCACTACTTCGGCTACAGACCCCTGGGGGTAT 240  
DB |||||  
QY 787 ACTTCAACGGGCGCCAGCAACAGCAACCACTACTTCGGCTACAGACCCCTGGGGGTAT 846  
DB |||||  
QY 241 TTTGATTTCAACAGATTTCCACTGCCACTTTTCAACAAGTGTGAGCGAGCTCATCAAC 300  
DB |||||  
QY 847 TTTGATTTCAACAGATTTCCACTGCCACTTTTCAACAAGTGTGAGCGAGCTCATCAAC 906  
DB |||||  
QY 301 AACAAATGGGGATTCGGGCGCCAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCGAG 360  
DB |||||  
QY 907 AACAAATGGGGATTCGGGCGCCAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCGAG 966  
DB |||||  
QY 361 GAGGTCACGACCAATGATGGCGTCAACACCATCGCTAATACTTACAGCAGCGTTCAA 420  
DB |||||  
QY 967 GAGGTCACGACCAATGATGGCGTCAACACCATCGCTAATACTTACAGCAGCGTTCAA 1026  
DB |||||  
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGTCTCGGCTCTGCGCACCGAGGCTGC 480  
DB |||||  
QY 1027 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGTCTCGGCTCTGCGCACCGAGGCTGC 1086  
DB |||||  
QY 481 CTCCCTCCGTTCCGGCGGAGCTGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 540  
DB |||||  
QY 1087 CTCCCTCCGTTCCGGCGGAGCTGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 1146  
DB |||||  
QY 541 AATGGCAGCAAGCGTGGGAGCTTCATCTTTACTGCTGGAATATTTCCCTTCTCAG 600  
DB |||||  
QY 1147 AATGGCAGCAAGCGTGGGAGCTTCATCTTTTACTGCTGGAATATTTCCCTTCTCAG 1206  
DB |||||  
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGGAGAGTGCCTTTCCAC 660  
DB |||||  
QY 1207 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGGAGAGTGCCTTTCCAC 1266  
DB |||||  
QY 661 AGCAGCTACCGCAAGCAGAGCCTGGACCGGCTGATGAATCTCTCATTCGACCAATAC 720  
DB |||||  
QY 1267 AGCAGCTACCGCAAGCAGAGCCTGGACCGGCTGATGAATCTCTCATTCGACCAATAC 1326  
DB |||||  
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAAGACTTTGCTG 780  
DB |||||  
QY 1327 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAAGACTTTGCTG 1386  
DB |||||  
QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAAACTGGCTACCTGGACCC 840  
DB |||||  
QY 1387 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTACGCCCAAAACTGGCTACCTGGACCC 1446  
DB |||||  
QY 841 TGTATTTCGCGAGCAGCGCGTTTCTAAACAAACAAACAGACAAACAAACAGCAATTTTACC 900  
DB |||||  
QY 1447 TGTATTTCGCGAGCAGCGCGTTTCTAAACAAACAAACAGACAAACAAACAGCAATTTTACC 1506  
DB |||||  
QY 901 TGGACTGGTGTCTCAAAATATATACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 960  
DB |||||  
QY 1507 TGGACTGGTGTCTCAAAATATATACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 1566  
DB |||||  
QY 961 GCTATGGCTTCACACAAAGACGAGCAAGTCTTTTCCCATGAGCGGTGTCATGATT 1020  
DB |||||  
QY 1567 GCTGTGGCTTCACACAAAGACGAGCAAGTCTTTTCCCATGAGCGGTGTCATGATT 1626  
DB |||||  
QY 1021 TTTGGAAAAAGAGAGCCCGGAGCTTCAAAACATCTGCAATGGACAAATGTCATGATACAGAC 1080  
DB |||||  
QY 1627 TTTGGAAAAAGAGAGCCCGGAGCTTCAAGCACTGTCATTTGGACAAATGTCATGATACAGAC 1686  
DB |||||

Search completed: November 28, 2005, 19:55:47

Job time : 754.542 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 12:51:27 ; Search time 5861.51 Seconds  
(without alignments)  
15555.198 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_2829\_4432  
Perfect score: 1604  
Sequence: 1 atggcttcagggtggcgc.....tacctaccgtccctgtga 1604

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.in.\*
- 3: gb.env.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pr.\*
- 9: gb.ro.\*
- 10: gb.sts.\*
- 11: gb.sy.\*
- 12: gb.un.\*
- 13: gb.vi.\*
- 14: gb.htg.\*
- 15: gb.pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1604	100.0	1605	6	BD242773 Adeno-ass
2	1604	100.0	1605	6	AR562505 Sequence
3	1604	100.0	1800	6	BD242772 Adeno-ass
4	1604	100.0	1800	6	AR562504 Sequence
5	1604	100.0	2211	6	BD242771 Adeno-ass
6	1604	100.0	2211	6	CS073491 Sequence
7	1604	100.0	2211	6	CS073614 Sequence
8	1604	100.0	2211	6	AR562503 Sequence
9	1604	100.0	4718	6	BD242766 Adeno-ass
10	1604	100.0	4718	6	AR562498 Sequence
11	1604	100.0	4718	6	AX753251 Sequence
12	1604	100.0	4718	13	AF063497 Adeno-ass
13	1604	100.0	7447	6	AR527492 Sequence
14	1600.8	99.8	2211	6	CS073493 Sequence
15	1600.8	99.8	2211	13	AY530607 Adeno-ass
16	1594.4	99.4	2214	6	CS073492 Sequence
17	1594.4	99.4	2214	13	AY530606 Adeno-ass
18	1592.8	99.3	2211	6	CS073490 Sequence

19	1592.8	99.3	2211	13	AY530611	Adeno-ass
20	1591.2	99.2	2211	6	CS073494	Sequence
21	1591.2	99.2	2211	13	AY530609	Adeno-ass
22	1512.8	94.3	2211	6	CS073615	Sequence
23	1512.8	94.3	4683	6	BD242775	Adeno-ass
24	1512.8	94.3	4683	6	AR562507	Sequence
25	1512.8	94.3	4683	13	AF028704	Adeno-ass
26	1104.6	68.9	2214	6	CS073452	Sequence
27	1104.6	68.9	2214	13	AY530567	Adeno-ass
28	1084	67.6	2214	6	CS073453	Sequence
29	1079	67.3	2214	6	CS073448	Sequence
30	1077.4	67.2	2214	6	CS073447	Sequence
31	1075.8	67.1	2217	6	CS073454	Sequence
32	1074.2	67.0	2214	6	CS073444	Sequence
33	1074.2	67.0	2214	6	CS073446	Sequence
34	1074.2	67.0	2214	13	AY530561	Adeno-ass
35	1072.6	66.9	2214	6	CS073451	Sequence
36	1071.6	66.8	2190	13	AY243000	Non-human
37	1071.6	66.8	3094	6	AX753267	Sequence
38	1071	66.8	2214	6	CS073445	Sequence
39	1071	66.8	2214	6	CS073449	Sequence
40	1071	66.8	2214	13	AY530568	Adeno-ass
41	1071	66.8	2214	13	AY530573	Adeno-ass
42	1070	66.7	2190	13	AY242998	Non-human
43	1070	66.7	2190	13	AY242999	Non-human
44	1070	66.7	2214	6	CS073625	Sequence
45	1070	66.7	3095	6	AX753268	Sequence

ALIGNMENTS

RESULT 1  
BD242773 1605 bp DNA linear PAT 17-JUL-2003  
LOCUS Adeno-associated virus serum type 1 nucleic acid sequence, vector  
DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector  
ACCESSION BD242773.1 GI:33052543  
KEYWORDS JP 2002529098-A/8.  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 1605)  
AUTHORS Wilson,J.M. and Xiao,W.  
TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector  
JOURNAL and host cell containing the same  
Patent: JP 2002529098-A 8 10-SEP-2002.  
COMMENT THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA  
OS AAV-1  
PN JP 2002529098-A/8  
PD 10-SEP-2002  
PF 02-NOV-1999 JP 2000581227  
PR 05-NOV-1998 US 60/107114  
PI JAMES M WILSON, WEIDONG XIAO  
PC C12N15/09,A61K31/711,A61P43/00,A61P43/00,C12N1/15,C12N1/19, PC  
C12N1/21,  
PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC  
Adeno-associated virus serum type 1 nucleic acid sequence, CC  
vector and host  
CC cell containing the same  
FH Key Location/Qualifiers  
FT CDS Location/Qualifiers  
1..1605  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

FEATURES  
source  
1..1605  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 100.0%; Score 1604; DB 6; Length 1605;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAAATAACGAAGGCGCCGACGGAGTGGGT 60  
DB 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAAATAACGAAGGCGCCGACGGAGTGGGT 60  
QY 61 AATGCTCAGGAAATTTGGCAATTTGGCAATTTCCATATGGCTGGGCGACAGAGTCAATCACCAC 120  
DB 61 AATGCTCAGGAAATTTGGCAATTTGGCAATTTCCATATGGCTGGGCGACAGAGTCAATCACCAC 120  
QY 121 AGCACCGCACCTGGGCGCTTGGCCACCTCAATTAACCACTCTTCAAGCAAAATCTCCAGT 180  
DB 121 AGCACCGCACCTGGGCGCTTGGCCACCTCAATTAACCACTCTTCAAGCAAAATCTCCAGT 180  
QY 181 GCTTCAACGGGGCCAGCAACCACTTTCACACGCTGACTGGCAGCGACTCATCAAC 240  
DB 181 GCTTCAACGGGGCCAGCAACCACTTTCACACGCTGACTGGCAGCGACTCATCAAC 240  
QY 241 TTTGATTTCAACAGATTTCCACTGCCACTTTTACACGCTGACTGGCAGCGACTCATCAAC 300  
DB 241 TTTGATTTCAACAGATTTCCACTGCCACTTTTACACGCTGACTGGCAGCGACTCATCAAC 300  
QY 301 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360  
DB 301 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360  
QY 361 GAGGTCAACGATGATGGGCTGACAACTGCTTAATTAACCTTACACGAGCGGTTCAA 420  
DB 361 GAGGTCAACGATGATGGGCTGACAACTGCTTAATTAACCTTACACGAGCGGTTCAA 420  
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTTCAGTACGCTCTCGGCTCGCGACACGAGGCTGC 480  
DB 421 GTCTTCTCGGACTCGGAGTACAGCTTTCAGTACGCTCTCGGCTCGCGACACGAGGCTGC 480  
QY 481 CTCCTCTCGGCTTCGCGGAGCGTGTTCATGATTTCCGCAATACGCGTACCTGACGCTCAAC 540  
DB 481 CTCCTCTCGGCTTCGCGGAGCGTGTTCATGATTTCCGCAATACGCGTACCTGACGCTCAAC 540  
QY 541 AATGGCAGCGCAAGCGTGGAGCGTTTCATGCTTTTACTGCTGGCAATATTTCCCTTCTCAG 600  
DB 541 AATGGCAGCGCAAGCGTGGAGCGTTTCATGCTTTTACTGCTGGCAATATTTCCCTTCTCAG 600  
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660  
DB 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660  
QY 661 AGCAGCTACGCGCACAGCAGAGCGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720  
DB 661 AGCAGCTACGCGCACAGCAGAGCGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720  
QY 721 CTGTATTACCTGAACAGAACTGAAATCAGTCGGGAAGTGCCTTCAACCAAGGACTTGGCTG 780  
DB 721 CTGTATTACCTGAACAGAACTGAAATCAGTCGGGAAGTGCCTTCAACCAAGGACTTGGCTG 780  
QY 781 TTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTTTCAGCCCAAACTGGCTTACCTGGACCC 840  
DB 781 TTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTTTCAGCCCAAACTGGCTTACCTGGACCC 840  
QY 841 TGTATTTCGGCAGCAGCGGCTTTCTTAAACCAAAACAGACAAACCAACGCAATTTTACC 900  
DB 841 TGTATTTCGGCAGCAGCGGCTTTCTTAAACCAAAACAGACAAACCAACGCAATTTTACC 900  
QY 901 TGAATGCTGCTTCAAAATATAACCTCAATGGGCGTGAATTCATCATCAACCTCGGACT 960  
DB 901 TGAATGCTGCTTCAAAATATAACCTCAATGGGCGTGAATTCATCATCAACCTCGGACT 960  
QY 961 GCTATGGCTTCACAAAGACGAGCAAGTCTTTCCTCCATGAGCGGTGTCATGATT 1020  
DB 961 GCTATGGCTTCACAAAGACGAGCAAGTCTTTCCTCCATGAGCGGTGTCATGATT 1020  
QY 1021 TTTGGAAAGAGAGCGCGGAGCTTCAAACTGCTGATTTGGCAATGTCTGATTTACAGAC 1080  
DB 1021 TTTGGAAAGAGAGCGCGGAGCTTCAAACTGCTGATTTGGCAATGTCTGATTTACAGAC 1080

QY 1081 GAAGAGGAAATTTAAAGCCACTAAACCTGCTGGCCACCGAAAGATTTGGGACCGTGGCAGTTC 1140  
DB 1081 GAAGAGGAAATTTAAAGCCACTAAACCTGCTGGCCACCGAAAGATTTGGGACCGTGGCAGTTC 1140  
QY 1141 AATTTTCCAGAGCAGCAGCAGCAGCAGCCTTCGACCGGAGATGTGCATGTCTATGGGAGCATTA 1200  
DB 1141 AATTTTCCAGAGCAGCAGCAGCAGCAGCCTTCGACCGGAGATGTGCATGTCTATGGGAGCATTA 1200  
QY 1201 CTTGCATGTGTGTGGCAAGATAGAGACCTGTACCTTCGAGGGTCCCATTTGGGCCAAAATT 1260  
DB 1201 CTTGCATGTGTGTGGCAAGATAGAGACCTGTACCTTCGAGGGTCCCATTTGGGCCAAAATT 1260  
QY 1261 CTTGCATGTGTGTGGCAAGATAGAGACCTGTACCTTCGAGGGTCCCATTTGGGCCAAAATT 1320  
DB 1261 CTTGCATGTGTGTGGCAAGATAGAGACCTGTACCTTCGAGGGTCCCATTTGGGCCAAAATT 1320  
QY 1321 CCGGCTCTCTCAGATCTCTCATCAAAACACGCGCTGTTCCTGCGAAATCTCCCGCGGAGTTT 1380  
DB 1321 CCGGCTCTCTCAGATCTCTCATCAAAACACGCGCTGTTCCTGCGAAATCTCCCGCGGAGTTT 1380  
QY 1381 TCAGCTCAAAAGTTTGTCTTCATTCATCAACCAATCTCAACGAGCAAGTGTGTGGA 1440  
DB 1381 TCAGCTCAAAAGTTTGTCTTCATTCATCAACCAATCTCAACGAGCAAGTGTGTGGA 1440  
QY 1441 ATTGAATGGGAGCTCGAGAAAGAAACAGCAGCGCTGGAATCCCGAAGTGCAGTACACA 1500  
DB 1441 ATTGAATGGGAGCTCGAGAAAGAAACAGCAGCGCTGGAATCCCGAAGTGCAGTACACA 1500  
QY 1501 TCCAAATATGCAAAATCTGCCAAACCTGTGATTTTACTGTGGCAACAATGGAATCTTATCT 1560  
DB 1501 TCCAAATATGCAAAATCTGCCAAACCTGTGATTTTACTGTGGCAACAATGGAATCTTATCT 1560  
QY 1561 GAGCCTCGCCCCCATTTGGCAACCGTTTACCTTACCGCTCCCTGTGA 1604  
DB 1561 GAGCCTCGCCCCCATTTGGCAACCGTTTACCTTACCGCTCCCTGTGA 1604

RESULT 2  
LOCUS AR562505 1605 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 16 from patent US 6759237.  
ACCESSION AR562505  
VERSION AR562505.1 GI:53976571  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1605)  
AUTHORS Wilson,J.M. and Xiao,W.  
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors and host cells containing same  
JOURNAL Patent: US 6759237-A 16 06-JUL-2004;  
The Trustees of the University of Pennsylvania; Philadelphia, PA  
FEATURES  
source  
1..1605  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 1604; DB 6; Length 1605;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAAATAACGAAGGCGCCGACGGAGTGGGT 60  
DB 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAAATAACGAAGGCGCCGACGGAGTGGGT 60  
QY 61 AATGCTCAGGAAATTTGGCAATTTGGCAATTTCCATATGGCTGGGCGACAGAGTCAATCACCAC 120  
DB 61 AATGCTCAGGAAATTTGGCAATTTGGCAATTTCCATATGGCTGGGCGACAGAGTCAATCACCAC 120  
QY 121 AGCACCGCACCTGGGCGCTTGGCCACCTCAATTAACCACTCTTCAAGCAAAATCTCCAGT 180  
DB 121 AGCACCGCACCTGGGCGCTTGGCCACCTCAATTAACCACTCTTCAAGCAAAATCTCCAGT 180  
QY 181 GCTTCAACGGGGCCAGCAACCACTTTCACACGCTGACTGGCAGCGACTCATCAAC 240  
DB 181 GCTTCAACGGGGCCAGCAACCACTTTCACACGCTGACTGGCAGCGACTCATCAAC 240  
QY 241 TTTGATTTCAACAGATTTCCACTGCCACTTTTACACGCTGACTGGCAGCGACTCATCAAC 300  
DB 241 TTTGATTTCAACAGATTTCCACTGCCACTTTTACACGCTGACTGGCAGCGACTCATCAAC 300  
QY 301 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360  
DB 301 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360  
QY 361 GAGGTCAACGATGATGGGCTGACAACTGCTTAATTAACCTTACACGAGCGGTTCAA 420  
DB 361 GAGGTCAACGATGATGGGCTGACAACTGCTTAATTAACCTTACACGAGCGGTTCAA 420  
QY 421 GTCTTCTCGGACTCGGAGTACAGCTTTCAGTACGCTCTCGGCTCGCGACACGAGGCTGC 480  
DB 421 GTCTTCTCGGACTCGGAGTACAGCTTTCAGTACGCTCTCGGCTCGCGACACGAGGCTGC 480  
QY 481 CTCCTCTCGGCTTCGCGGAGCGTGTTCATGATTTCCGCAATACGCGTACCTGACGCTCAAC 540  
DB 481 CTCCTCTCGGCTTCGCGGAGCGTGTTCATGATTTCCGCAATACGCGTACCTGACGCTCAAC 540  
QY 541 AATGGCAGCGCAAGCGTGGAGCGTTTCATGCTTTTACTGCTGGCAATATTTCCCTTCTCAG 600  
DB 541 AATGGCAGCGCAAGCGTGGAGCGTTTCATGCTTTTACTGCTGGCAATATTTCCCTTCTCAG 600  
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660  
DB 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660  
QY 661 AGCAGCTACGCGCACAGCAGAGCGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720  
DB 661 AGCAGCTACGCGCACAGCAGAGCGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720  
QY 721 CTGTATTACCTGAACAGAACTGAAATCAGTCGGGAAGTGCCTTCAACCAAGGACTTGGCTG 780  
DB 721 CTGTATTACCTGAACAGAACTGAAATCAGTCGGGAAGTGCCTTCAACCAAGGACTTGGCTG 780  
QY 781 TTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTTTCAGCCCAAACTGGCTTACCTGGACCC 840  
DB 781 TTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTTTCAGCCCAAACTGGCTTACCTGGACCC 840  
QY 841 TGTATTTCGGCAGCAGCGGCTTTCTTAAACCAAAACAGACAAACCAACGCAATTTTACC 900  
DB 841 TGTATTTCGGCAGCAGCGGCTTTCTTAAACCAAAACAGACAAACCAACGCAATTTTACC 900  
QY 901 TGAATGCTGCTTCAAAATATAACCTCAATGGGCGTGAATTCATCATCAACCTCGGACT 960  
DB 901 TGAATGCTGCTTCAAAATATAACCTCAATGGGCGTGAATTCATCATCAACCTCGGACT 960  
QY 961 GCTATGGCTTCACAAAGACGAGCAAGTCTTTCCTCCATGAGCGGTGTCATGATT 1020  
DB 961 GCTATGGCTTCACAAAGACGAGCAAGTCTTTCCTCCATGAGCGGTGTCATGATT 1020  
QY 1021 TTTGGAAAGAGAGCGCGGAGCTTCAAACTGCTGATTTGGCAATGTCTGATTTACAGAC 1080  
DB 1021 TTTGGAAAGAGAGCGCGGAGCTTCAAACTGCTGATTTGGCAATGTCTGATTTACAGAC 1080

Db	121	AGCACCGCACCTGGGGCTTGGCCACCTACAAATAACCACTCTACAAAGCAAAATCTCCAGT	180
Qy	181	GCCTTCAACGGGGCCAGCAACAGCAACCACTACTCTCGGTACAGCAACCCCTGGGGGTAT	240
Db	181	GCCTTCAACGGGGCCAGCAACAGCAACCACTACTCTCGGTACAGCAACCCCTGGGGGTAT	240
Qy	241	TTTTGATTTCAACAGATTCACCTGCCACTTTTACCAAGCTGACTGGCAGGACTCATCAAC	300
Db	241	TTTTGATTTCAACAGATTCACCTGCCACTTTTACCAAGCTGACTGGCAGGACTCATCAAC	300
Qy	301	AACAAATGGGGATTCGCGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCTCAAGTCAAG	360
Db	301	AACAAATGGGGATTCGCGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCTCAAGTCAAG	360
Qy	361	GAGGTACAGCAAGATGATGCGTCAACCAACATCGCTAATAACCTTACCAGCAGGTCAA	420
Db	361	GAGGTACAGCAAGATGATGCGTCAACCAACATCGCTAATAACCTTACCAGCAGGTCAA	420
Qy	421	GTCTTCTCGACTCGGAGTACAGCTTCGGTACGTCTCGGCTCTGGCACCAGGGCTGC	480
Db	421	GTCTTCTCGACTCGGAGTACAGCTTCGGTACGTCTCGGCTCTGGCACCAGGGCTGC	480
Qy	481	CTCCCTCCGTTCCGGCGGACGTGTTTCATGATTCGCAATACGGCTACCTGACGCCTCAAC	540
Db	481	CTCCCTCCGTTCCGGCGGACGTGTTTCATGATTCGCAATACGGCTACCTGACGCCTCAAC	540
Qy	541	AATGGCAGCCAAGCCGTGGGACGTTTCATCTCTTTTATCTGCCTGGAAATATTTCCCTTCTCAG	600
Db	541	AA TGGCAGCCAAGCCGTGGGACGTTTCATCTCTTTTATCTGCCTGGAAATATTTCCCTTCTCAG	600
Qy	601	ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGAGAGTGCCTTCCAC	660
Db	601	ATGCTGAGAACGGGCAACAACTTTACCTTCAGCTACACCTTTGAGAGAGTGCCTTCCAC	660
Qy	661	AGCAGCTACGGCAGCAGCAGCGCTGGACGGCTGATGAATCCTCTCATCGACCAATAC	720
Db	661	AGCAGCTACGGCAGCAGCAGCGCTGGACGGCTGATGAATCCTCTCATCGACCAATAC	720
Qy	721	CTGTATTACCTGTAACAGAACTCAAAAATCAGTCGGGAGTCCCAAAAACAGGACTTGGCTG	780
Db	721	CTGTATTACCTGTAACAGAACTCAAAAATCAGTCGGGAGTCCCAAAAACAGGACTTGGCTG	780
Qy	781	TTTAGCCGTGGGTCTCAGCTGGCATGTCTGTTTACGCCCAAAAACCTGGCTACCTGGACCC	840
Db	781	TTTAGCCGTGGGTCTCAGCTGGCATGTCTGTTTACGCCCAAAAACCTGGCTACCTGGACCC	840
Qy	841	TGTTATCGGCAGCAGCGGTTTCTAAACAARAACAGACACACACACACACATTTTACC	900
Db	841	TGTTATCGGCAGCAGCGGTTTCTAAACAARAACAGACACACACACACACATTTTACC	900
Qy	901	TGGACTGCTGCTTCAAAATATAACTCAATGGGGGTGAATCCATCATCAACCCCTGGCAGCT	960
Db	901	TGGACTGCTGCTTCAAAATATAACTCAATGGGGGTGAATCCATCATCAACCCCTGGCAGCT	960
Qy	961	GCTATGGCCTCACACAAAGACGACGAAGACAAAGTTCTTTCCCATGAGCGGTGCATGATT	1020
Db	961	GCTATGGCCTCACACAAAGACGACGAAGACAAAGTTCTTTCCCATGAGCGGTGCATGATT	1020
Qy	1021	TTTGGAAAAGAGAGCGCGGAGCTTCAAAACNCTGCATTTGACATGTATGATTACAGAC	1080
Db	1021	TTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGTCATTTGGACATGTATGATTACAGAC	1080
Qy	1081	GAAGAGGAAATTAAGCCACTTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC	1140
Db	1081	GAAGAGGAAATTAAGCCACTTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC	1140
Qy	1141	AATTTCCAGACGACGACACAGACCCCTCGACCGGAGATGTGATGCTATGGGACATTA	1200
Db	1141	AATTTCCAGACGACGACACAGACCCCTCGACCGGAGATGTGATGCTATGGGACATTA	1200
Qy	1201	CCTGGCATTGGTGTGGCAAGATAGAGACGTGTACTCTGAGGGTCCCATTTGGGCCAAAATT	1260
Db	1201	CCTGGCATTGGTGTGGCAAGATAGAGACGTGTACTCTGAGGGTCCCATTTGGGCCAAAATT	1260

Qy	1361	CCTCACACAGATGGACACTTTTCACCCCGTCTCCTCTTATGGCGGCTTTGGACTCAAGAAC	1320
Db	1361	CCTCACACAGATGGACACTTTTCACCCCGTCTCCTCTTATGGCGGCTTTGGACTCAAGAAC	1320
Qy	1321	CCGGCTCCTCAGATCCCTCATCAAAAAACACGCCCTGTTCTTCGCAATCCTCCGGCGGAGTTT	1380
Db	1321	CCGGCTCCTCAGATCCCTCATCAAAAAACACGCCCTGTTCTTCGCAATCCTCCGGCGGAGTTT	1380
Qy	1381	TCAGCTACAAAGTTTGTTTCATTCATCACCCCAATATCTCCACAGACAAGTGAGTGTGGAA	1440
Db	1381	TCAGCTACAAAGTTTGTTTCATTCATCACCCCAATATCTCCACAGACAAGTGAGTGTGGAA	1440
Qy	1441	ATTGAATGGGAGCTGCAGAAAGAAAACAGACGGCTGGAAATCCCGAAGTGCAGTACACA	1500
Db	1441	ATTGAATGGGAGCTGCAGAAAGAAAACAGACGGCTGGAAATCCCGAAGTGCAGTACACA	1500
Qy	1501	TCCAAATTATGCAAAATCTGCCCAACGTTTGATTTTACTTGGACAACAATGGACTTTTACT	1560
Db	1501	TCCAAATTATGCAAAATCTGCCCAACGTTTGATTTTACTTGGACAACAATGGACTTTTACT	1560
Qy	1561	GAGCCTCGCCCAATGGACCCGTTACTTACCCGTCCTCCCTGTGA	1604
Db	1561	GAGCCTCGCCCAATGGACCCGTTACTTACCCGTCCTCCCTGTGA	1604
RESULT 3	BD242772	1800 bp DNA linear	PAT 17-JUL-2003
LOCUS	BD242772	Adeno-associated virus serum type 1 nucleic acid sequence, vector	
DEFINITION		and host cell containing the same.	
ACCESSION	BD242772		
VERSION	BD242772.1	GI:33052542	
KEYWORDS	JP 2002529098-A/7.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1800)		
AUTHORS	Wilson, J.M. and Xiao, W.		
TITLE	Adeno-associated virus serum type 1 nucleic acid sequence, vector		
JOURNAL	Patent: JP 2002529098-A 7 10-SEP-2002;		
COMMENT	THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA		
	OS AAV-1		
	PN JP 2002529098-A/7		
	PD 10-SEP-2002		
	PF 02-NOV-1999 JP 2000581227		
	PR 05-NOV-1998 US 60/107114		
	PI JAMES M WILSON, WEIDONG XIAO		
	PC C12N15/09, A61K31/71, A61K48/00, A61P43/00, C12N1/15, C12N1/19, P06C12N1/21,		
	PC C12N5/10, C12P21/02, A61K35/76, C12N15/00, C12N5/00 CC		
	Adeno-associated virus serum type 1 nucleic acid sequence, CC		
	vector and host		
	CC cell containing the same		
FEATURES	Location/Qualifiers		
source	1..1800		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
ORIGIN			
Query Match	100.0%;	Score 1604;	DB 6; Length 1800;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1604;	Conservative	0; Mismatches	0; Indels 0; Gaps 0
Qy	1	ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGACGAGTGGGT	60
Db	196	ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGACGAGTGGGT	255
Qy	61	AATGCCTCAGGAAATGGCATTTGGCAATGGCTGGGGGACAGAGTCAATCACCAC	120

[illegible]

Db	1336	AAATTTCCAGACGACGACACAGACCTTCGACCGGAGATGTGTGATCTATGGGAGACATTA	1399
Qy	1201	CCTGGCATGGTGTGGCAAGATAGAGACGTGTACTTCGACGGGTCCCATTATGGGCGCAAAATTT	1260
Db	1396	CCTGGCATGGTGTGGCAAGATAGAGACGTGTACTTCGACGGGTCCCATTATGGGCGCAAAATTT	1455
Qy	1261	CCTCACACAGATGGACACTTTTACCCGGTCTCCTCTTTATGGGCGGCTTTGGACTCAAGAAC	1320
Db	1456	CCTCACACAGATGGACACTTTTACCCGGTCTCCTCTTTATGGGCGGCTTTGGACTCAAGAAC	1515
Qy	1321	CCGCGTCTCCAGATCCTCATCAAAAACAGCGCTTCTTCGCGAATCCTCCGGCGGAGTTT	1380
Db	1516	CCGCGTCTCCAGATCCTCATCAAAAACAGCGCTTCTTCGCGAATCCTCCGGCGGAGTTT	1575
Qy	1381	TCAGCTACAAAGTTTGCTTCATTTCATCACCCCAATACTCCACAGGACAAGTGAAGTGTGGAA	1440
Db	1576	TCAGCTACAAAGTTTGCTTCATTTCATCACCCCAATACTCCACAGGACAAGTGAAGTGTGGAA	1635
Qy	1441	ATTGAATGGGAGCTGCAGAAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA	1500
Db	1636	ATTGAATGGGAGCTGCAGAAAAGAAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA	1695
Qy	1501	TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGCACAAACAATGACTTTTATACT	1560
Db	1696	TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGCACAAACAATGACTTTTATACT	1755
Qy	1561	GAGCCTCGCCCCATTGGCACCCGTTACCTTACCGTCCCTCTGTA	1604
Db	1756	GAGCCTCGCCCCATTGGCACCCGTTACCTTACCGTCCCTCTGTA	1799
RESULT 4			
AR562504			
LOCUS	AR562504	1800 bp	DNA linear PAT 08-OCT-2004
DEFINITION	Sequence 14 from patent US 6759237.		
ACCESSION	AR562504		
VERSION	AR562504.1	GI:53976570	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1800)		
AUTHORS	Wilson,J.M. and Xiao,W.		
TITLE	Adeno-associated virus serotype 1 nucleic acid sequences, vectors and host cells containing same		
JOURNAL	Patent: US 6759237-A 14 06-JUL-2004;		
FEATURES	The Trustees of the University of Pennsylvania; Philadelphia, PA		
source	Location/Qualifiers		
	1..1800		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Query Match	100.0%;	Score 1604;	DB 6; Length 1800;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1604;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	ATGCGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCGCAGCGAGTGGGT	60
Db	196	ATGCGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCGCAGCGAGTGGGT	255
Qy	61	AATGCGCTCAGGAATTTGGCATTCGGATTCCAGATGGCTGGGCGCACAGAGTCATCCACC	120
Db	256	AATGCGCTCAGGAATTTGGCATTCGGATTCCAGATGGCTGGGCGCACAGAGTCATCCACC	315
Qy	121	AGCACCCGCACCTGGGCGCTTGGCCACCTCAATAAACCACTCTTACAAGCAAAATCTCCAGT	180
Db	316	AGCACCCGCACCTGGGCGCTTGGCCACCTCAATAAACCACTCTTACAAGCAAAATCTCCAGT	375
Qy	181	GCTTCAACGGGGCCAGCAACAGCAACCACTACTTGGGCTACAGCACCCCTCGGGGGTAT	240
Db	376	GCTTCAACGGGGCCAGCAACAGCAACCACTACTTGGGCTACAGCACCCCTCGGGGGTAT	435



```
QY 181 GCTTCAACGGGGCCAGCAACGACACCACTACTTTCGGCTACAGCACACCCCTGGGGTAT 240
Db 787 GCTTCAACGGGGCCAGCAACGACACCACTACTTTCGGCTACAGCACACCCCTGGGGTAT 846
QY 241 TTTGATTTCAACAGATTCCGACATGCGCACTTTTCAACACGTGACTGGCAGGAGCTCATCAAC 300
Db 847 TTTGATTTCAACAGATTCCGACATGCGCACTTTTCAACACGTGACTGGCAGGAGCTCATCAAC 906
QY 301 AACAAATGGGGATTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360
Db 907 AACAAATGGGGATTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 966
QY 361 GAGGTCACGACGAAATGATGGCGTCACAAACATCGCTAATAAATCTTACCGACGAGTTCAA 420
Db 967 GAGGTCACGACGAAATGATGGCGTCACAAACATCGCTAATAAATCTTACCGACGAGTTCAA 1026
QY 421 GTCTTCTCGGACTCGGAGTACAGACTTTCGGTACGCTCGGCTCTGCGCACACGAGGCTGC 480
Db 1027 GTCTTCTCGGACTCGGAGTACAGACTTTCGGTACGCTCGGCTCTGCGCACACGAGGCTGC 1086
QY 481 CTCCTCTCGTTCGGCGGACGTTTCATGATTCGGCAATACGGCTACCTGACGCTCAAC 540
Db 1087 CTCCTCTCGTTCGGCGGACGTTTCATGATTCGGCAATACGGCTACCTGACGCTCAAC 1146
QY 541 AATGGCAGCAACCGCTGGGAGCTTCATCTTTTACTGCTGGAAATATTTCCCTTCTCAG 600
Db 1147 AATGGCAGCAACCGCTGGGAGCTTCATCTTTTACTGCTGGAAATATTTCCCTTCTCAG 1206
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660
Db 1207 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 1266
QY 661 AGCAGCTACGCGCACAGCAGCCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 720
Db 1267 AGCAGCTACGCGCACAGCAGCCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 1326
QY 721 CTGTATTACTGACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAAGAGACTTGCTG 780
Db 1327 CTGTATTACTGACAGAACTCAAAATCAGTCGGGAAGTGCCCAAAAGAGACTTGCTG 1386
QY 781 TTTAGCCGTGGGTCTCCAGTCGATGTCTGTTCAGCCCAAAACCTGGCTACTCGACCC 840
Db 1387 TTTAGCCGTGGGTCTCCAGTCGATGTCTGTTCAGCCCAAAACCTGGCTACTCGACCC 1446
QY 841 TGTATTGCGCAGCAGCGCGTTTCTAAAAAAGAGCAAGCAACAAAGCAATTTTACC 900
Db 1447 TGTATTGCGCAGCAGCGCGTTTCTAAAAAAGAGCAAGCAACAAAGCAATTTTACC 1506
QY 901 TGGACTGGTGCTTCAAAATATTAACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 960
Db 1507 TGGACTGGTGCTTCAAAATATTAACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 1566
QY 961 GCTATGGCTTCACAAAGACGACGAAGCAAGTCTTTTCCCATGAGCGGTGTCATGATT 1020
Db 1567 GCTATGGCTTCACAAAGACGACGAAGCAAGTCTTTTCCCATGAGCGGTGTCATGATT 1626
QY 1021 TTTGGAAAAAGAGCGCCGGAGGTTTCAAACTGTGCAATTTGGACAAATGATGATTAAGAC 1080
Db 1627 TTTGGAAAAAGAGCGCCGGAGGTTTCAAACTGTGCAATTTGGACAAATGATGATTAAGAC 1686
QY 1081 GAAGAGGAAATTAAGGCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGGAGTC 1140
Db 1687 GAAGAGGAAATTAAGGCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGGAGTC 1746
QY 1141 AATTTCCAGAGCAGCAGACAGACCTTCGACCGGAGATGTGCTATCGGAGCATTA 1200
Db 1747 AATTTCCAGAGCAGCAGACAGACCTTCGACCGGAGATGTGCTATCGGAGCATTA 1806
QY 1201 CTTGGCATGGTGGCAAGATAGAGACGTTGATCTTCAGGCTCCCATTTTGGGCCAAAAAT 1260
Db 1807 CTTGGCATGGTGGCAAGATAGAGACGTTGATCTTCAGGCTCCCATTTTGGGCCAAAAAT 1866
```

```
QY 1261 CCTCACAGATGGACACTTTTACCGCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1320
Db 1867 CCTCACAGATGGACACTTTTACCGCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1926
QY 1321 CGGCTCTCAGATCTCTCATCAAAAAACACGCTGTTCCTGCGAATCTCTCCGGCGAGTTT 1380
Db 1927 CGGCTCTCAGATCTCTCATCAAAAAACACGCTGTTCCTGCGAATCTCTCCGGCGAGTTT 1986
QY 1381 TCAGTCAAAAGTTGGTTCATTCATACCCCAATCTCCACAGGACAAAGTGTGTGAA 1440
Db 1987 TCAGTCAAAAGTTGGTTCATTCATACCCCAATCTCCACAGGACAAAGTGTGTGAA 2046
QY 1441 ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGAAAGTCAGTACACA 1500
Db 2047 ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGAAAGTCAGTACACA 2106
QY 1501 TCCAATTATGCAAAATCTGCCAAGCTTGAATTTTACTGTGGACAAACAATGGACTTTAT 1560
Db 2107 TCCAATTATGCAAAATCTGCCAAGCTTGAATTTTACTGTGGACAAACAATGGACTTTAT 2166
QY 1561 GAGCTCTGCCCATTTGGCACCCGTTACCTTACCGTCCCTGTGA 1604
Db 2167 GAGCTCTGCCCATTTGGCACCCGTTACCTTACCGTCCCTGTGA 2210

RESULT 6
CS073491          2211 bp      DNA      linear      PAT 05-MAY-2005
LOCUS             Sequence 79 from Patent WO200503321.
DEFINITION
ACCESSION         CS073491
VERSION           CS073491.1 GI:63090470
KEYWORDS
SOURCE             .
ORGANISM           unidentifed
                   unidentifed
                   unclassified.
REFERENCE          1
AUTHORS            Wilson,J.M., Gao,G., Alvira,M.R. and Vandenbergh,L.H.
TITLE              Adeno-associated virus (aav) clades, sequences, vectors containing
                   same, and uses therefor
JOURNAL            Patent: WO 200503321-A 79 14-APR-2005;
                   The Trustees of the University of Pennsylvania (US)
FEATURES
source             1..2211
                   /organism="unidentifed"
                   /mol_type="unassigned DNA"
                   /db_xref="taxon:32644"
                   /note="new AAV serotype, clone rh.71"

ORIGIN
Query Match       100.0%; Score 1604; DB 6; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGCGCAGACAAATAACGAAGGCCGCGAGTGGGT 60
Db 607 ATGGCTTCAGGCGGTGGCGCACCAATGCGCAGACAAATAACGAAGGCCGCGAGTGGGT 666
QY 61 AATGCTTCAGGAAATTTGGCATTTGGCATTTCCATGGCTGGGCGACAGAGTCAATCAACC 120
Db 667 AATGCTTCAGGAAATTTGGCATTTGGCATTTCCATGGCTGGGCGACAGAGTCAATCAACC 726
QY 121 AGCACCCGCACTGGGCGCTTGGCCACCTTACATTAACCACTCTACAGCAAAATCTCCAGT 180
Db 727 AGCACCCGCACTGGGCGCTTGGCCACCTTGGCCACCTTACATTAACCACTCTACAGCAAAATCTCCAGT 786
QY 181 GCTTCAACGGGGCCAGCAACGACAAACCACTACTTTCGGCTACAGCACACCCCTGGGGTAT 240
Db 787 GCTTCAACGGGGCCAGCAACGACAAACCACTACTTTCGGCTACAGCACACCCCTGGGGTAT 846
QY 241 TTTGATTTCAACAGATTCCACTGCGCACTTTTTCACCACTGACTGGCAGGACTCATCAAC 300
Db 847 TTTGATTTCAACAGATTCCACTGCGCACTTTTTCACCACTGACTGGCAGGACTCATCAAC 906
```





Db 1027 GTCTTCGGACTCGGATGACGAGTTCGGTACGCTTCGGTCTCGGACACGAGGCTGC 1086  
Qy 481 CTCCTCCGTTCCCGCGGACGTTGATGATTCGCAATACGGCTACCTGACGCTCAAC 540  
Db 1087 CTCCTCCGTTCCCGCGGACGTTGATGATTCGCAATACGGCTACCTGACGCTCAAC 1146  
Qy 541 AATGGCAGCAGCCGTTGGAGCTTCATCTCTTTACTGCTGGAATATTTCCCTTCTCAG 600  
Db 1147 AATGGCAGCAGCCGTTGGAGCTTCATCTCTTTACTGCTGGAATATTTCCCTTCTCAG 1206  
Qy 601 ATGCTGAGAACGGGCAACAACTTTACTCTTACGTCACCTTTTGAGGAAGTGCCTTTCCAC 660  
Db 1207 ATGCTGAGAACGGGCAACAACTTTACTCTTACGTCACCTTTTGAGGAAGTGCCTTTCCAC 1266  
Qy 661 AGCAGCTACCGCACAGCAGCCTGGACCGGCTGATGAATCTCTCTCATCGACCAATAC 720  
Db 1267 AGCAGCTACCGCACAGCAGCCTGGACCGGCTGATGAATCTCTCTCATCGACCAATAC 1326  
Qy 721 CTGTATTACCTGAAAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAAGAAAGTGTG 780  
Db 1327 CTGTATTACCTGAAAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAAGAAAGTGTG 1386  
Qy 781 TTTAGCCGTGGTCTCCAGCTGGCATGCTGTTTCAGCCCAAAACTGGCTACTCGACCC 840  
Db 1387 TTTAGCCGTGGTCTCCAGCTGGCATGCTGTTTCAGCCCAAAACTGGCTACTCGACCC 1446  
Qy 841 TGTATTCCGACGAGCGCGTTTCTAAACAAAAACAGACACAAACAGCAATTTTACC 900  
Db 1447 TGTATTCCGACGAGCGCGTTTCTAAACAAAAACAGACACAAACAGCAATTTTACC 1506  
Qy 901 TGGACTGTGCTTCAAAATATTAACCTCAATGGGCGTGAATTCATTCATCAACCCCTGGCACT 960  
Db 1507 TGGACTGTGCTTCAAAATATTAACCTCAATGGGCGTGAATTCATTCATCAACCCCTGGCACT 1566  
Qy 961 GCTATGGCTCTCACAAAGACGACGAAGCAAGTCTTTCCTCATGACGCGTGTCAATGTT 1020  
Db 1567 GCTATGGCTCTCACAAAGACGACGAAGCAAGTCTTTCCTCATGACGCGTGTCAATGTT 1626  
Qy 1021 TTTGGAAAAAGAGCGCCGGAGCTTCAAAACACTGCAATGGACAAATGTCATGATTAACAGAC 1080  
Db 1627 TTTGGAAAAAGAGCGCCGGAGCTTCAAAACACTGCAATGGACAAATGTCATGATTAACAGAC 1686  
Qy 1081 GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140  
Db 1687 GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1746  
Qy 1141 AATTTCCAGACGACGACAGACCTCGACCGGAGATGTGCATGCTATGGAGCATTA 1200  
Db 1747 AATTTCCAGACGACGACAGACCTCGACCGGAGATGTGCATGCTATGGAGCATTA 1806  
Qy 1201 CTTGGCATGTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCAATTTGGGCCAAAAT 1260  
Db 1807 CTTGGCATGTGTGGCAAGATAGAGACGTGTACCTGACGGGTCCCAATTTGGGCCAAAAT 1866  
Qy 1261 CTTCAACAGATGGACACTTTTCAACCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1320  
Db 1867 CTTCAACAGATGGACACTTTTCAACCGCTCTCTCTTATGGCGGCTTTGGACTCAAGAAC 1926  
Qy 1321 CGCCCTCTCAGATCTCATCAAAAACAGCGCTTCTTCCGGAATCTCCGGCGGAGTTT 1380  
Db 1927 CGCCCTCTCAGATCTCATCAAAAACAGCGCTTCTTCCGGAATCTCCGGCGGAGTTT 1986  
Qy 1381 TCAGCTACAAAGTTTGTCTTCTTATTCATCAACCAATCTCCACAGGACAAAGTGTGGAA 1440  
Db 1987 TCAGCTACAAAGTTTGTCTTCTTATTCATCAACCAATCTCCACAGGACAAAGTGTGGAA 2046  
Qy 1441 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCAAGTGCAGTACACA 1500  
Db 2047 ATTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCAAGTGCAGTACACA 2106  
Qy 1501 TCCAAATTAACAATACTCGCAACGTTGATTTTACTGTGGACAAACATGACTTTTATCT 1560

Db 2107 TCCAAATTAACAATACTCGCAACGTTGATTTTACTGTGGACAAATGAGACTTTATCT 2166  
Qy 1561 GAGCTCTGCCCATTTGGCACCCTGTTACCTTACCGTCCCTCTGTA 1604  
Db 2167 GAGCTCTGCCCATTTGGCACCCTGTTACCTTACCGTCCCTCTGTA 2210  
RESULT 8  
AR562503 LOCUS 2211 bp DNA linear PAT 08-OCT-2004  
DEFINITION Sequence 12 from patent US 6759237.  
ACCESSION AR562503  
VERSION AR562503.1 GI:53976569  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2211)  
AUTHORS Wilson,J.M. and Xiao,W.  
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors and host cells containing same  
JOURNAL Patent: US 6759237-A 12 06-JUL-2004;  
The Trustees of the University of Pennsylvania; Philadelphia, PA  
FEATURES  
source 1..2211  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 100.0%; Score 1604; DB 6; Length 2211;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGCAGGAGTGGGT 60  
Db 607 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGCAGGAGTGGGT 666  
Qy 61 AATGCTCAGGAAATTTGGCATTTGGATTCACATGGCTGGGGGACAGAGTATCACACC 120  
Db 667 AATGCTCAGGAAATTTGGCATTTGGCATTTCCATGGCTGGGGGACAGAGTATCACACC 726  
Qy 121 AGCACCCGACCTGGGCTTGGCCACCTACCAATTAACCACTCTCAACGCAAAATCTCCAGT 180  
Db 727 AGCACCCGACCTGGGCTTGGCCACCTACCAATTAACCACTCTCAACGCAAAATCTCCAGT 786  
Qy 181 GCTTCAACGGGGGCGCAGCAACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT 240  
Db 787 GCTTCAACGGGGGCGCAGCAACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT 846  
Qy 241 TTTGATTTCAACAGATTTCCACTGCGCACTTTTCAACGCTGACTGGCAGGACTCATCAAC 300  
Db 847 TTTGATTTCAACAGATTTCCACTGCGCACTTTTCAACGCTGACTGGCAGGACTCATCAAC 906  
Qy 301 AACAAATTCGGGATTCGGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAAGTCAAG 360  
Db 907 AACAAATTCGGGATTCGGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAAGTCAAG 966  
Qy 361 GAGGTACGACGAAATGATGGGCTCAACACCATCGCTTAATAACCTTTACCAGCAACGGTTCAA 420  
Db 967 GAGGTACGACGAAATGATGGGCTCAACACCATCGCTTAATAACCTTTACCAGCAACGGTTCAA 1026  
Qy 421 GTCTTCTCGGACTCGGAGTACAGGTTCCGTCAGTTCCTGGCTTCGGGACACAGGCTGC 480  
Db 1027 GTCTTCTCGGACTCGGAGTACAGGTTTCCTGCTACGCTCTCGGCTTCGCGCACAGGCTGC 1086  
Qy 481 CTCCTCTCGGCTCCCGGGGAGCGTGTTCATGATTCGCAATACGGCTACTCGTACGCTCAAC 540  
Db 1087 CTCCTCTCGGCTCCCGGGGAGCGTGTTCATGATTCGCAATACGGCTACTCGTACGCTCAAC 1146  
Qy 541 AATGGCAGCAGCCGTTGGAGCTTCATCTCTTTTACTGCTGGAATATTTCCCTTCTCAG 600  
Db 1147 AATGGCAGCAGCCGTTGGAGCTTCATCTCTTTTACTGCTGGAATATTTCCCTTCTCAG 1206

QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660  
Db 1207 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 1266  
QY 661 AGCAGCTACGGCAGACGACGAGCCTTGGACCGGCTGATGANTCTCTCATCGACCAATAC 720  
Db 1267 AGCAGCTACGGCAGACGACGAGCCTTGGACCGGCTGATGANTCTCTCATCGACCAATAC 1326  
QY 721 CTGTATTACTGTAACAGAACTCAAAATCAGTCCGGAAAGTGCCTTGGAAAGCAAGCACTTGCTG 780  
Db 1327 CTGTATTACTGTAACAGAACTCAAAATCAGTCCGGAAAGTGCCTTGGAAAGCAAGCACTTGCTG 1386  
QY 781 TTTAGCGTGGGTCTCCAGCTGGCATGTCTGTTTTCAGCCCAAAACTGGCTACTCGACCC 840  
Db 1387 TTTAGCGTGGGTCTCCAGCTGGCATGTCTGTTTTCAGCCCAAAACTGGCTACTCGACCC 1446  
QY 841 TGTATTCCGAGCAGCGCGTCTTCAAAACAAAACAGACACAAACAGCAATTTTACC 900  
Db 1447 TGTATTCCGAGCAGCGCGTCTTCAAAACAAAACAGACACAAACAGCAATTTTACC 1506  
QY 901 TGGACTGGTCTTCAAAATATTAACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 960  
Db 1507 TGGACTGGTCTTCAAAATATTAACCTCAATGGGCGTGAATCCATCATCAACCTGGCACT 1566  
QY 961 GCTATGGCTTCACAAAGACGACGAAGCAAGTCTTTTCCCATGAGCGGTGTGATGTT 1020  
Db 1567 GCTATGGCTTCACAAAGACGACGAAGCAAGTCTTTTCCCATGAGCGGTGTGATGTT 1626  
QY 1021 TTTGGAAAGAGAGCGCGGAGCTTCAACACATGCAATGGACAAATGTCATGATTACAGAC 1080  
Db 1627 TTTGGAAAGAGAGCGCGGAGCTTCAACACATGCAATGGACAAATGTCATGATTACAGAC 1686  
QY 1081 GAAGAGGAATTAAGCCACTTAACCTCTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140  
Db 1687 GAAGAGGAATTAAGCCACTTAACCTCTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1746  
QY 1141 AATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTGATGCTATGGGACGATTA 1200  
Db 1747 AATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTGATGCTATGGGACGATTA 1806  
QY 1201 CTTGGCATGTGTGGCAAGATAGAGAGCTGTA CTTGAGGGTCCCATTTTGGGCCAAATTT 1260  
Db 1807 CTTGGCATGTGTGGCAAGATAGAGAGCTGTA CTTGAGGGTCCCATTTTGGGCCAAATTT 1866  
QY 1261 CCTCAGACAGATGGACACTTTTCAACCGCTCTCTTTATGGCGGCTTTGGACCTCAAGAAC 1320  
Db 1867 CCTCAGACAGATGGACACTTTTCAACCGCTCTCTTTATGGCGGCTTTGGACCTCAAGAAC 1926  
QY 1321 CCGCCTCCTCAGATCCTCATCAAAAAACACGCTCTGTTCTCGGAATCCTCGGCGGAGTTT 1380  
Db 1927 CCGCCTCCTCAGATCCTCATCAAAAAACACGCTCTGTTCTCGGAATCCTCGGCGGAGTTT 1986  
QY 1381 TCAGCTACAAAGTTTGTTCATCATCAACCAATATCTCCAAGCAAGTGAAGTGGAA 1440  
Db 1987 TCAGCTACAAAGTTTGTTCATCATCAACCAATATCTCCAAGCAAGTGAAGTGGAA 2046  
QY 1441 ATTGAATGGAGCTGCAGAAAGAAACAGACGCTGGGAATCCCGAAGTGCAGTACACA 1500  
Db 2047 ATTGAATGGAGCTGCAGAAAGAAACAGACGCTGGGAATCCCGAAGTGCAGTACACA 2106  
QY 1501 TCCAAATTATGCAAAATCTGCAACGTTTGAATTTTACTGTGGACAAATGGAATTTTACT 1560  
Db 2107 TCCAAATTATGCAAAATCTGCAACGTTTGAATTTTACTGTGGACAAATGGAATTTTACT 2166  
QY 1561 GAGCCTCGCCCAATTTGACACCGTTTACTTTACCGGTCCCTCTGA 1604  
Db 2167 GAGCCTCGCCCAATTTGACACCGTTTACTTTACCGGTCCCTCTGA 2210

RESULT 9  
BD242766  
LOCUS  
DEFINITION BD242766 4718 bp DNA linear PAT 17-JUL-2003  
Adeno-associated virus serum type 1 nucleic acid sequence, vector

and host cell containing the same.  
BD242766  
VERSION BD242766.1 GI:33052536  
KEYWORDS JP 2002529098-A/1.  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 4718)  
AUTHORS Wilson,J.M. and Xiao,W.  
TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector  
and host cell containing the same  
JOURNAL Patent: JP 2002529098-A 1 10-SEP-2002;  
COMMENT THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA  
OS AAV-1  
PN JP 2002529098-A/1  
PD 10-SEP-2002  
PP 02-NOV-1999 JP 2000581227  
PR 05-NOV-1998 US 60/107114  
PI JAMES M WILSON,WEIDONG XIAO  
PC C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19, PC  
C12N1/21,  
PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC  
Adeno-associated virus serum type 1 nucleic acid sequence, CC  
vector and host  
CC cell containing the same  
FH Key Location/Qualifiers  
FT CDS (335)..(2206)  
FT CDS (2223)..(4430).  
FEATURES  
source 1..4718  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"  
ORIGIN  
Query Match 100.0%; Score 1604; DB 6; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCGCAGCGAGTGGGT 60  
Db 2829 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCGCAGCGAGTGGGT 2888  
QY 61 AATGCCTCAGGAATTTGGCATTGCGATTCCCATGCTGGGCGACAGAGTCATCACCAC 120  
Db 2889 AATGCCTCAGGAATTTGGCATTGCGATTCCCATGCTGGGCGACAGAGTCATCACCAC 2948  
QY 121 AGACCCGACCTGGGCGCTTGGCCACCTCAATTAACCACTCTTACAAGAAATCTCCAGT 180  
Db 2949 AGACCCGACCTGGGCGCTTGGCCACCTCAATTAACCACTCTTACAAGAAATCTCCAGT 3008  
QY 181 GCTTCAACGGGGCGCAGCAACCACTACTTCGGCTACAGACCCCTCTGGGGGTAT 240  
Db 3009 GCTTCAACGGGGCGCAGCAACCACTACTTCGGCTACAGACCCCTCTGGGGGTAT 3068  
QY 241 TTTGATTTCAACAGATTCCCATGCTCCCACTTTTACCACGCTGACTGGCAGCGACTCATCAAC 300  
Db 3069 TTTGATTTCAACAGATTCCCATGCTCCCACTTTTACCACGCTGACTGGCAGCGACTCATCAAC 3128  
QY 301 AACAAATTTGGGGAATTCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAGTCAAG 360  
Db 3129 AACAAATTTGGGGAATTCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAGTCAAG 3188  
QY 361 GAGGTCAACGACGAATGATGGCGTCACAACCATCGCTAAATAAATCTTACCAGCAGGTTCAA 420  
Db 3189 GAGGTCAACGACGAATGATGGCGTCACAACCATCGCTAAATAAATCTTACCAGCAGGTTCAA 3248  
QY 421 GTCTTCTCGGACTCGGAGTACAGACTTCCGTACGTCTCTCGGCTCTGCGCACCCAGGGCTGC 480  
Db 3249 GTCTTCTCGGACTCGGAGTACAGACTTCCGTACGTCTCTCGGCTCTGCGCACCCAGGGCTGC 3308  
QY 481 CTCCTCTCGGTTCCCGCGGACAGTGTTCATGATTCGCGAATACGGCTACCTGAGCGCTCAAC 540  
|||||

3309 CTCCCTCGTTCCCGCGGACGTTTCATGATTCGCGAATACGGCTACCTGACGCTCAAC 3368  
QY 541 AATGGCAGCAGCGTGGACGTTTCATCTCTTTTACTGCTGGAATATTTCCCTTCTAG 600  
Db 3369 AATGGCAGCAGCGTGGACGTTTCATCTCTTTTACTGCTGGAATATTTCCCTTCTAG 3428  
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGGAGAGTGCCTTCCAC 660  
Db 3429 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGGAGAGTGCCTTCCAC 3488  
QY 661 AGCAGCTACGCGACAGCAGCGCTGGACCGGCTGATGAATCTCTCTCATCGACCAATAC 720  
Db 3489 AGCAGCTACGCGACAGCAGCGCTGGACCGGCTGATGAATCTCTCATCGACCAATAC 3548  
QY 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAAAGAGACTTGTG 780  
Db 3549 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTGCCTCAAAAAGAGACTTGTG 3608  
QY 781 TTTAGCGGTGGTCTCCAGCTGGCATGCTGTTTCAGCCCAAAAACCTGGCTACCTGACCC 840  
Db 3609 TTTAGCGGTGGTCTCCAGCTGGCATGCTGTTTCAGCCCAAAAACCTGGCTACCTGACCC 3668  
QY 841 TGTATTCCGACAGCGCGTCTTAAACAAAAACAGACAAACAAACAGCAATTTTACC 900  
Db 3669 TGTATTCCGACAGCGCGTCTTAAACAAAAACAGACAAACAAACAGCAATTTTACC 3728  
QY 901 TGGACTGGTCTTCAAAATATTAACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 960  
Db 3729 TGGACTGGTCTTCAAAATATTAACCTCAATGGCGTGAATCCATCATCAACCTGGCACT 3788  
QY 961 GCTATGGCTCACAAAAACAGACGAGCAAGTCTTTCCCATAGCGGTGTCATGATT 1020  
Db 3789 GCTATGGCTCACAAAAACAGACGAGCAAGTCTTTCCCATAGCGGTGTCATGATT 3848  
QY 1021 TTTGAAAAAGAGCGCCGGAGCTTCAAAACATGCAATTTGGACAAATGTCATGATTACAGAC 1080  
Db 3849 TTTGAAAAAGAGCGCCGGAGCTTCAAAACATGCAATTTGGACAAATGTCATGATTACAGAC 3908  
QY 1081 GAAGAGAAATTAAGCCACTAACCTGTGGCGACCGAAAGATTTTGGACCGTGGAGTTC 1140  
Db 3909 GAAGAGAAATTAAGCCACTAACCTGTGGCGACCGAAAGATTTTGGACCGTGGAGTTC 3968  
QY 1141 AATTTCCAGACGACGACAGACCTGCGACCGGAGATGTGCTGCTATGGAGCATTA 1200  
Db 3969 AATTTCCAGACGACGACAGACCTGCGACCGGAGATGTGCTGCTATGGAGCATTA 4028  
QY 1201 CTGGCATGGTGTGGCAAGATAGACGCTGACCTGACGGTCCCAATTTGGGCAAAAT 1260  
Db 4029 CTGGCATGGTGTGGCAAGATAGACGCTGACCTGACGGTCCCAATTTGGGCAAAAT 4088  
QY 1261 CCTCACAGATGGACACTTTCACCGTCTCTCTTATGGGGGCTTTGGACTCAAGAAC 1320  
Db 4089 CCTCACAGATGGACACTTTCACCGTCTCTCTTATGGGGGCTTTGGACTCAAGAAC 4148  
QY 1321 CGCGCTCTCAGATCTCATCAAAAACAGCTGTTCTTCGCGAATCTCCCGCGGAGTTT 1380  
Db 4149 CGCGCTCTCAGATCTCATCAAAAACAGCTGTTCTTCGCGAATCTCCCGCGGAGTTT 4208  
QY 1381 TCAGCTACAAAGTTTGTCTTATTTCATCACCCCAATCTCCACAGGACAAAGTGTGGAA 1440  
Db 4209 TCAGCTACAAAGTTTGTCTTATTTCATCACCCCAATCTCCACAGGACAAAGTGTGGAA 4268  
QY 1441 ATTGAATGGAGTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 1500  
Db 4269 ATTGAATGGAGTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAGTACACA 4328  
QY 1501 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAAGTGTGTTTACT 1560  
Db 4329 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAAGTGTGTTTACT 4388  
QY 1561 GAGCCTCGCCCCATTTGGCACCCGTTTACCTTTACCGTCCCTGTGA 1604  
Db 4389 GAGCCTCGCCCCATTTGGCACCCGTTTACCTTTACCGTCCCTGTGA 4432

RESULT 10  
AR562498  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
ORIGIN  
Query Match  
Best Local Similarity  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGACGGAGTGGGT 60  
Db 2829 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGACGGAGTGGGT 2888  
QY 61 AATGCTCAGGAAATTTGGCATTTCCATGCTGGTGGCGCAGAGTCAATCACCACC 120  
Db 2889 AATGCTCAGGAAATTTGGCATTTCCATGCTGGTGGCGCAGAGTCAATCACCACC 2948  
QY 121 AGCACCCGACCTGGGCTTGGCCACCTACAAATAACCACTCTACAAGCAAAATCTCCAGT 180  
Db 2949 AGCACCCGACCTGGGCTTGGCCACCTACAAATAACCACTCTACAAGCAAAATCTCCAGT 3008  
QY 181 GCTTCAACGGGGCGCGACCAACGACCAACCACTACTTCGGCTACAGCACCCCTCGGGGTTAT 240  
Db 3009 GCTTCAACGGGGCGCGACCAACGACCAACCACTACTTCGGCTACAGCACCCCTCGGGGTTAT 3068  
QY 241 TTTGATTTCAACAGATTTCACTGGCAGCTTTTCCACGAGTACTGGCAGGAGTCAATCAAC 300  
Db 3069 TTTGATTTCAACAGATTTCACTGGCAGCTTTTCCACGAGTACTGGCAGGAGTCAATCAAC 3128  
QY 301 AACAATTTGGGATTTCCGGCCCAAGAGACTCAAACTTCAAACTCTTCAAACTCTTCAAACTCAAGTCAAG 360  
Db 3129 AACAATTTGGGATTTCCGGCCCAAGAGACTCAAACTTCAAACTCTTCAAACTCTTCAAACTCAAGTCAAG 3188  
QY 361 GAGGTCACGACGAATGATGGGCTCAAAACCATCGCTTAATAAATTTACCAGACGGTTCAA 420  
Db 3189 GAGGTCACGACGAATGATGGGCTCAAAACCATCGCTTAATAAATTTACCAGACGGTTCAA 3248  
QY 421 GTCTTCTGGACTCGGAGTACAGCTTCCGAGTCTCTGGCTCTGGCGACGAGGCTGC 480  
Db 3249 GTCTTCTGGACTCGGAGTACAGCTTCCGAGTCTCTGGCTCTGGCGACGAGGCTGC 3308  
QY 481 CTCCCTCGTTCCTGGGAGCTGTTTCATGATTCGCAATACGGTACTCGACGCTCAAC 540  
Db 3309 CTCCCTCGTTCCTGGGAGCTGTTTCATGATTCGCAATACGGTACTCGACGCTCAAC 3368  
QY 541 AATGGCAGCAAGCGGTGGAGCTTCACTCTTTTACTGCTGGAATATTTCCCTTCTCAG 600  
Db 3369 AATGGCAGCAAGCGGTGGAGCTTCACTCTTTTACTGCTGGAATATTTCCCTTCTCAG 3428  
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGGAGAGTGCCTTCCAC 660  
Db 3429 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGGAGAGTGCCTTCCAC 3488  
QY 661 AGCAGCTACGCGACAGCAGCGCTGGACCGGCTGATGAATCTCTCTCATCGACCAATAC 720

AR562498  
Sequence 1 from patent US 6759237.  
linear PAT 08-OCT-2004  
AR562498  
AR562498.1 GI:53976564  
Unknown.  
Unknow.  
Unclassified.  
1 (bases 1 to 4718)  
Wilson,J.M. and Xiao,W.  
Adeno-associated virus serotype 1 nucleic acid sequences, vectors  
and host cells containing same  
Patent: US 6759237-A 1 06-JUL-2004;  
The Trustees of the University of Pennsylvania; Philadelphia, PA  
Location/Qualifiers  
1. 4718  
/organism="unknown"  
/mol\_type="genomic DNA"

Db 3489 AGCAGCTACGCGCACAGCGCCTGGAGCCGGCTGATGAATCCTCTCATCGACCAATAC 3548  
Qy 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAGGACTTGCTG 780  
Db 3549 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAGGACTTGCTG 3608  
Qy 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCCAAAACCTGGCTACTCGGACCC 840  
Db 3609 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCCAAAACCTGGCTACTCGGACCC 3668  
Qy 841 TGTATTCCGCGAGCAGCGCGTCTTAAACAAAACAGACAAACAAACAGCAATTTTACC 900  
Db 3669 TGTATTCCGCGAGCAGCGCGTCTTAAACAAAACAGACAAACAAACAGCAATTTTACC 3728  
Qy 901 TGGACTGTGTCTTCAAAATATACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 960  
Db 3729 TGGACTGTGTCTTCAAAATATACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 3788  
Qy 961 GCTATGGCTTCACACAAAGACGACGAAGCAAGTCTTTCCCATGAGCGGTGTCTATGATT 1020  
Db 3789 GCTATGGCTTCACACAAAGACGACGAAGCAAGTCTTTCCCATGAGCGGTGTCTATGATT 3848  
Qy 1021 TTTGGAAAAGAGAGCGCCGGAGCTTCAAAACATGCAATTTGGACAATGTCTATGATTACAGAC 1080  
Db 3849 TTTGGAAAAGAGAGCGCCGGAGCTTCAAAACATGCAATTTGGACAATGTCTATGATTACAGAC 3908  
Qy 1081 GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140  
Db 3909 GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 3968  
Qy 1141 AATTTCCAGAGCAGCAGCAGACCTCTGGACCGGAGATGTGCTATGGAGCATTTA 1200  
Db 3969 AATTTCCAGAGCAGCAGCAGACCTCTGGACCGGAGATGTGCTATGGAGCATTTA 4028  
Qy 1201 CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCATTTGGGCCAAAAT 4088  
Db 4029 CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCATTTGGGCCAAAAT 4148  
Qy 1261 CCTCACAGATGGACACTTTCACCCGTCTCTTATGGCGGCTTTGGACTCAAGAAC 1320  
Db 4089 CCTCACAGATGGACACTTTCACCCGTCTCTTATGGCGGCTTTGGACTCAAGAAC 4148  
Qy 1321 CGCCCTCTCAGATCTCTCAATAAAACACGCTGTTCTTCGCAATCTCTCGGCGGAGTTT 1380  
Db 4149 CGCCCTCTCAGATCTCTCAATAAAACACGCTGTTCTTCGCAATCTCTCGGCGGAGTTT 4208  
Qy 1381 TCAGCTACAAAGTTTGTCTCATTCATCAACCCAACTCTCCACAGGACAAGTGTGGAA 1440  
Db 4209 TCAGCTACAAAGTTTGTCTCATTCATCAACCCAACTCTCCACAGGACAAGTGTGGAA 4268  
Qy 1441 ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCCGAAGTGCAGTACACA 1500  
Db 4269 ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCCGAAGTGCAGTACACA 4328  
Qy 1501 TCCAAATATGCAAAATCTGCAACGTTGATTTTACTGTGGACAACTGGACTTTTACT 1560  
Db 4329 TCCAAATATGCAAAATCTGCAACGTTGATTTTACTGTGGACAACTGGACTTTTACT 4388  
Qy 1561 GAGCCTCGCCCCATTTGGACCCGTTACCTTTACCCGTCCTCTGTA 1604  
Db 4389 GAGCCTCGCCCCATTTGGACCCGTTACCTTTACCCGTCCTCTGTA 4432

RESULT 11  
AX753251  
LOCUS AX753251 4718 bp DNA linear PAT 23-JUN-2003  
DEFINITION Sequence 6 from Patent EP1310571.  
ACCESSION AX753251  
VERSION AX753251.1 GI:32166108  
KEYWORDS Adeno-associated virus 1  
SOURCE Adeno-associated virus 1  
ORGANISM Adeno-associated virus 1

Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.  
1  
REFERENCE  
AUTHORS Gao, G., Wilson, J.M. and Alvira, M.  
TITLE A method of detecting and/or identifying adeno-associated virus (AVV) sequences and isolating novel sequences identified thereby  
JOURNAL Patent: EP 1310571-A 6 14-MAY-2003;  
The Trustees of the University of Pennsylvania (US)  
FEATURES  
Location/Qualifiers  
1..4718  
/organism="Adeno-associated virus 1"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:85106"  
ORIGIN  
Query Match 100.0%; Score 1604; DB 6; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGCTTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCCGACGAGTGGGT 60  
Db 2829 ATGGCTTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCCGACGAGTGGGT 2888  
Qy 61 AATGCTCTCAGGAATTTGGCATTTGCCATTCCACATGCTGGCGACAGAGTTCATCACACC 120  
Db 2889 AATGCTCTCAGGAATTTGGCATTTGCCATTCCACATGCTGGCGACAGAGTTCATCACACC 2948  
Qy 121 AGCACCCGACCTGGGCTTGGCCACCTACATAAACCACCTCTTCAAGCAAAATCTCCAGT 180  
Db 2949 AGCACCCGACCTGGGCTTGGCCACCTACATAAACCACCTCTTCAAGCAAAATCTCCAGT 3008  
Qy 181 GCTTCAACGGGGCGCAGCAACCAACCACTCTTCGGCTACAGCAACCCCTTGGGGGTAT 240  
Db 3009 GCTTCAACGGGGCGCAGCAACCAACCACTCTTCGGCTACAGCAACCCCTTGGGGGTAT 3068  
Qy 241 TTTGATTTTCAACAGATTTCCACTGCTCCACTTTTCCACAGCTGACTGGCAGGACTCATCAAC 300  
Db 3069 TTTGATTTTCAACAGATTTCCACTGCTCCACTTTTCCACAGCTGACTGGCAGGACTCATCAAC 3128  
Qy 301 AAACAAATTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 360  
Db 3129 AAACAAATTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAAGTCAAG 3188  
Qy 361 GAGGTCAACGCAATGATGGCGTCAACACATCGCTTAATACCTTACAGACGCGTTCAA 420  
Db 3189 GAGGTCAACGCAATGATGGCGTCAACACATCGCTTAATACCTTACAGACGCGTTCAA 3248  
Qy 421 GTCTTCTCGACTCGGAGTACAGACTTCCGTCCTCGGCTCTCGGACGAGGCTGC 480  
Db 3249 GTCTTCTCGACTCGGAGTACAGACTTCCGTCCTCGGCTCTCGGACGAGGCTGC 3308  
Qy 481 CTCCCTTCGTTCCCGCGGACGCTGTTTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 540  
Db 3309 CTCCCTTCGTTCCCGCGGACGCTGTTTCATGATTTCCGCAATACGGCTACCTGACGCTCAAC 3368  
Qy 541 AATGGCAGCAAGCGCTGGGACGTTTCATCTCTTTACTGCTGGAATATTTCCTTCTTCAG 600  
Db 3369 AATGGCAGCAAGCGCTGGGACGTTTCATCTCTTTACTGCTGGAATATTTCCTTCTTCAG 3428  
Qy 601 ATGCTGAGAACGGGCAACAACTTTTACTCTTACGCTACACTTTGAGGAGTGCCTTTCCAC 660  
Db 3429 ATGCTGAGAACGGGCAACAACTTTTACTCTTACGCTACACTTTGAGGAGTGCCTTTCCAC 3488  
Qy 661 AGCAGCTACGCGCACAGCCAGAGCCTGGACCGGCTGATGAATCCTCTCTCGACCAATATAC 720  
Db 3489 AGCAGCTACGCGCACAGCCAGAGCCTGGACCGGCTGATGAATCCTCTCTCGACCAATATAC 3548  
Qy 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAGGACTTGCTG 780  
Db 3549 CTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAGGACTTGCTG 3608  
Qy 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCCAAAACCTGGCTACTCGGACCC 840  
Db 3609 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCCAAAACCTGGCTACTCGGACCC 3668



QY	841	TGTTATCGGACGACGCGTCTTAAACAAACAAACAGACAAACAAACAGCAATTTTACC	900
Db	3669	TGTTATCGGACGACGCGTCTTAAACAAACAAACAGACAAACAAACAGCAATTTTACC	3728
QY	901	TGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATCATCATCAACCTGGCACT	960
Db	3729	TGGACTGGTGTCTTCAAAATATAACCTCAATGGCGGTGAATCATCATCAACCTGGCACT	3788
QY	961	GCTATGGCTCTACACAAAGACGACGAAGACAAGTCTTTCCCATGAGCGGTGTCAATGATT	1020
Db	3789	GCTATGGCTCTACACAAAGACGACGAAGACAAGTCTTTCCCATGAGCGGTGTCAATGATT	3848
QY	1021	TTTGGAAAAGAGAGCGCCGGAGCTTCAACACTGCAATGGCAATGTCATGATTAACAGAC	1080
Db	3849	TTTGGAAAAGAGAGCGCCGGAGCTTCAACACTGCAATGGCAATGTCATGATTAACAGAC	3908
QY	1081	GAAGAGGAATTAAGCCCACTAAACCTGTGGCCACCGAAGAGATTTGGGACCGTGGCAGTC	1140
Db	3909	GAAGAGGAATTAAGCCCACTAAACCTGTGGCCACCGAAGAGATTTGGGACCGTGGCAGTC	3968
QY	1141	AATTTCCAGACGACGACGACAGACCTGCGACCGGAGATGTGCAATGATGGAGCATTA	1200
Db	3969	AATTTCCAGACGACGACGACAGACCTGCGACCGGAGATGTGCAATGATGGAGCATTA	4028
QY	1201	CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCGAGGTCCCATTTTGGGCCAAATTT	1260
Db	4029	CCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCGAGGTCCCATTTTGGGCCAAATTT	4088
QY	1261	CCTCACACAGATGGACACTTTTACCCGCTCTCTCTTATGGCGGCTTGGACTCAAGAAC	1320
Db	4089	CCTCACACAGATGGACACTTTTACCCGCTCTCTCTTATGGCGGCTTGGACTCAAGAAC	4148
QY	1321	CGCCTCTCAGATCCTCATCAAAAACACGCCCTGTTCTCGCAATCCTCCGGCGGAGTTT	1380
Db	4149	CGCCTCTCAGATCCTCATCAAAAACACGCCCTGTTCTCGCAATCCTCCGGCGGAGTTT	4208
QY	1381	TCAGCTACAAAGTTTGTCTTATTCATCACCATACTCACAGGACGAAGTGTGGA	1440
Db	4209	TCAGCTACAAAGTTTGTCTTATTCATCACCATACTCACAGGACGAAGTGTGGA	4268
QY	1441	ATTGAATGGAGCTGCAGAAAGAAAACAGCAGCGCTGGAATCCCGAAGTGCAGTACACA	1500
Db	4269	ATTGAATGGAGCTGCAGAAAGAAAACAGCAGCGCTGGAATCCCGAAGTGCAGTACACA	4328
QY	1501	TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGACTTTTATCT	1560
Db	4329	TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGACTTTTATCT	4388
QY	1561	GAGCCTCGCCCATTTGGCACCCGTTACCTTACCGTCCCTCTGTA	1604
Db	4389	GAGCCTCGCCCATTTGGCACCCGTTACCTTACCGTCCCTCTGTA	4432
RESULT 12			
AF063497			
LOCUS	AF063497	4718 bp	DNA linear VRL 27-APR-1999
DEFINITION	Adeno-associated virus 1, complete genome.		
ACCESSION	AF063497		
VERSION	AF063497.1 GI:4689096		
KEYWORDS			
SOURCE	Adeno-associated virus 1		
ORGANISM	Adeno-associated virus 1		
REFERENCE	Xiao, W., Chirmule, N., Berta, S.C., McCullough, B., Gao, G. and Wilson, J.M.		
AUTHORS	Xiao, W., Chirmule, N., Berta, S.C., McCullough, B., Gao, G. and Wilson, J.M.		
TITLE	Gene therapy vectors based on adeno-associated virus type 1		
JOURNAL	J. Virol. 73 (5), 3994-4003 (1999)		
PUBMED	10196295		
REFERENCE	2 (bases 1 to 4718)		
AUTHORS	Xiao, W. and Wilson, J.M.		
TITLE	Direct Submission		

JOURNAL	Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601 Spruce Street, Philadelphia, PA 19104, USA		
FEATURES			
source	1. 4718		
	/organism="Adeno-associated virus 1"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:85106"		
	/note="AAV1"		
	335..2206		
CDS			
	/codon_start=1		
	/product="nonstructural protein"		
	/protein_id="AAD27758.1"		
	/db_xref="GI:4689098"		
	/translation="MPGEYEVIVKPSLDLDELHLCISDSFVSVAEKWELPPDSMDMLNLEQALQAEKLDLQVLRVSKAPALFVQEKGESYFHLHILVELTGVKS		
	MVLRFSLQIRDLKLVQITIRGIEPTLPNMFVTKTRNAGGNGKVVDCYIPNYLLPK		
	TOPELQAWTMEVYISACLNLAKRKLVAQHLTHVSQTSQNKENLNPNSDAPVIRS		
	TKSAPDYLGVGPAPPADIKTNRIYRIELUNGVEPAVAGSVFLGWAOKRFGKNTIWLFG		
	PATGKTINIAEIAHAYPPGCVNNTNENFFPNCVDKVIWHEGKMTAKVESA		
	ILGSKVRVDQKQSSAQIDTPVIIVTNNCAVIDGNTSTTFHQQLDQWMPKFL		
	TRRLHDFGKVTQGEVFEFRWAQDHVTEVAHEFYRKGANKRPAPDADKSEPKRA		
	CPSVADPSTDAEGPVDADRYQNKCSRHALQMLFPCKTCERMNFIICFTGHT		
	RDCSECFPGVSESQPVVRKTRTKLCAIHLILGRAPEIACACDLVNVLDLDCVSEQ		
	2223..4433		
	/codon_start=1		
	/product="capsid protein"		
	/protein_id="AAD27757.1"		
	/db_xref="GI:4689097"		
	/translation="MAADGYLPDMLNLSIGIREWMDLKPQAPKPKANQKQDDGRG		
	LVLPQKYLGPNGLDKGPVNAADAALHDKAYDQOLKAGDNPYLRYNHADEFOE		
	RLQEDTSFGNLRGAVFOAKRVLPLEPLGVEGAKTAPGKKRPVQSGFQEPSSSGIG		
	KIQQPAKKRLNFGTQDSVPDQPLGEPPTAAVGPITTMASGGGAPMADNNEGA		
	DVGNSAGNHCDSITMLGDRVITSTRWALPTYNHLYKQISSASTGASNDVHTFY		
	STPMGYDFNRFHCFSPRDLNNNNWFRPKLNFLEKFNIOQVEVTTNDVHTTIA		
	NNLTSVQVSDSEYOLPYLGSAGHGLPPPADVFMIPQVGYLTLANGSOAVGRSS		
	PYCLEYFPQMLRTGNNTFSYTFEEVPHSHSAHSQSLDRMLNPLIDQYLYLARTQ		
	NQSQAQNKLFLSRGSPAGMSVQPKWLPQPCYRQQRVSKTKTNNNSNFTWTGASK		
	YNLNRESIINPGTAMASHKODEDKFPFQSGVMIPGKESAGASNTALONWITDSEI		
	KATNPVATERETVAVNFQSSDTPDGVHAMGALPGKWMQDRDVIQGPVIAWKIPIH		
	TDGHPHPSPLMGSGFLGPPQIILIKNTVPVNPAPFAFSATKFSFIFQYSGQVSE		
	IEWELQKENSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL"		
ORIGIN			
Query Match	100.0%;	Score 1604;	DB 13; Length 4718;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1604;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAAGAGGCGCGCAGCGAGTGGGT	60
Db	2829	ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAAGAGGCGCGCAGCGAGTGGGT	2888
QY	61	AATGCTCAGGAAATTTGGCATTTCCATGCTGGCTGGGCGCAGAGTCAATCCACC	120
Db	2889	AATGCTCAGGAAATTTGGCATTTCCATGCTGGCTGGGCGCAGAGTCAATCCACC	2948
QY	121	AGCACCGCACCTGGCGCTTGCACCACTTCCACCACTGAGTGGGCGCAGAGTCAATCC	180
Db	2949	AGCACCGCACCTGGCGCTTGCACCACTTCCACCACTGAGTGGGCGCAGAGTCAATCC	3008
QY	181	GCTTCAACGGGGCGCAGCAACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT	240
Db	3009	GCTTCAACGGGGCGCAGCAACCACTACTTTCGGCTACAGCACCCCTGGGGGTAT	3068
QY	241	TTTGATTTTCAACAGATTCACATGCGCACTTTTCCACCACTGAGTGGGCGCAGTCAATCA	300
Db	3069	TTTGATTTTCAACAGATTCACATGCGCACTTTTCCACCACTGAGTGGGCGCAGTCAATCA	3128
QY	301	AACAATTTGGGATTTCCGGCCCAAGAGACTCAACTCTTCAAACTCTTCAAGTCAAGTCAAG	360
Db	3129	AACAATTTGGGATTTCCGGCCCAAGAGACTCAACTCTTCAAACTCTTCAAGTCAAGTCAAG	3188
QY	361	GAGGTACACGACGATGATGGCGTCACACCATCGCTAATACTTACCAGCAGCGTTCAT	420

Db 3189 GAGGTACACGAGAAATGATGGCGTCAACACATCGCTAAATAACCTTACCAGCACGGTTCAA 3248  
QY 421 GTCTTCTCGGACTCGGAGTACAGACTTCCGTAGTCTCGGCTCTCGGCACACAGGGCTGC 480  
Db 3249 GTCTTCTCGGACTCGGAGTACAGACTTCCGTAGTCTCGGCTCTCGGCACACAGGGCTGC 3308  
QY 481 CTCCCTCCGTTCCCGCGGACAGTGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 540  
Db 3309 CTCCCTCCGTTCCCGCGGAGCGTGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 3368  
QY 541 AATGGCAGCAAGCCGCTGGAGCGTTTCATCTTTTACTGTGCTGGAATATTTCCCTTCTCAG 600  
Db 3369 AATGGCAGCAAGCCGCTGGAGCGTTTCATCTTTTACTGTGCTGGAATATTTCCCTTCTCAG 3428  
QY 601 ATGCTGAGAACGGGCAACACTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTCCAC 660  
Db 3429 ATGCTGAGAACGGGCAACACTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTCCAC 3488  
QY 661 AGCAGCTACGCGCACAGCCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720  
Db 3489 AGCAGCTACGCGCACAGCCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 3548  
QY 721 CTGTATTACTGAACAGAACTCAAAATCAATCGGAAAGTGCCTCAAAACAGAACTTGTCTG 780  
Db 3549 CTGTATTACTGAACAGAACTCAAAATCAATCGGAAAGTGCCTCAAAACAGAACTTGTCTG 3608  
QY 781 TTTAGCCGTGGTCTCCAGCTGCGCATGTCTGTTACGCCCAAAACCTGGCTACTCGACCC 840  
Db 3609 TTTAGCCGTGGTCTCCAGCTGCGCATGTCTGTTACGCCCAAAACCTGGCTACTCGACCC 3668  
QY 841 TGTATTCCGCGCAGCGCGTCTTCAAAACAAACAGACAAACAAACAGCAATTTTACC 900  
Db 3669 TGTATTCCGCGCAGCGCGTCTTCAAAACAAACAGACAAACAAACAGCAATTTTACC 3728  
QY 901 TGGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 960  
Db 3729 TGGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCCCTGGCACT 3788  
QY 961 GCTATGGCTCACACAAAGCAGAGCAAGTCTTCTTCCCATGAGCGGTGCATGATT 1020  
Db 3789 GCTATGGCTCACACAAAGCAGAGCAAGTCTTCTTCCCATGAGCGGTGCATGATT 3848  
QY 1021 TTTGGAAGAGAGCGCGGAGCTTCAACACATGCAATGGCAATGTCATGATTAACAGAC 1080  
Db 3849 TTTGGAAGAGAGCGCGGAGCTTCAACACATGCAATGGCAATGTCATGATTAACAGAC 3908  
QY 1081 GAAGAGGAAATTAAGCCACTAACCCCTGTGGCCACCCGAAGAATTTGGGACCGTGGCAGTC 1140  
Db 3909 GAAGAGGAAATTAAGCCACTAACCCCTGTGGCCACCCGAAGAATTTGGGACCGTGGCAGTC 3968  
QY 1141 AATTTCCAGAGCAGCAGACAGACCCCTGCGACCGGAGATGTGATGCTATGGGAGCATTA 1200  
Db 3969 AATTTCCAGAGCAGCAGACAGACCCCTGCGACCGGAGATGTGATGCTATGGGAGCATTA 4028  
QY 1201 CCTGGCATGTGTGGCAAGATAGAGAGCTGCTACCTGAGGCTCCCATTTGGGCCAAAATT 1260  
Db 4029 CCTGGCATGTGTGGCAAGATAGAGAGCTGCTACCTGAGGCTCCCATTTGGGCCAAAATT 4088  
QY 1261 CCTCACACAGATGGACACTTTTACCCCGTCTCTTATTTGGGCGGCTTTGGACTCAAGAAC 1320  
Db 4089 CCTCACACAGATGGACACTTTTACCCCGTCTCTTATTTGGGCGGCTTTGGACTCAAGAAC 4148  
QY 1321 CCGCCTCTCAGATCTCATCAAAACACAGCCTGTTCTCGGAAATCTCCGGGGAGTTT 1380  
Db 4149 CCGCCTCTCAGATCTCATCAAAACACAGCCTGTTCTCGGAAATCTCCGGGGAGTTT 4208  
QY 1381 TCAGCTACAAAGTTTGTCTTATTCATCAACCAATCTCCACAGCAAGTGAAGTGGAA 1440  
Db 4209 TCAGCTACAAAGTTTGTCTTATTCATCAACCAATCTCCACAGCAAGTGAAGTGGAA 4268  
QY 1441 ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAGTACACA 1500

Db 4269 ATTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGAAAGTGCAGTACACA 4328  
QY 1501 TCCTAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTTACT 1560  
Db 4329 TCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTTACT 4388  
QY 1561 GAGCCTCGCCCATTTGGACACCGGTTACCTTACCGTCCCTCTGTA 1604  
Db 4389 GAGCCTCGCCCATTTGGACACCGGTTACCTTACCGTCCCTCTGTA 4432

RESULT 13  
AR527492 LOCUS  
DEFINITION Sequence 11 from patent US 6723551.  
ACCESSION AR527492  
VERSION AR527492.1 GI:53914590  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 7447)  
AUTHORS Kotin, R.M., Urabe, M. and Ding, C.-T.  
TITLE Production of adeno-associated virus in insect cells  
JOURNAL Patent: US 6723551-A 11 20-APR-2004;  
The United States of America as represented by the Department of  
Health and Human Services; Washington, DC

FEATURES  
source  
1..7447  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 1604; DB 6; Length 7447;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1604; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGTGGGT 60  
Db 5227 ATGGCTTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGAGTGGGT 5286  
QY 61 AATGCTCAGGAATTTGGCATTTCCGATTCACATGCTGGCGCAGAGTCAACACC 120  
Db 5287 AATGCTCAGGAATTTGGCATTTCCGATTCACATGCTGGCGCAGAGTCAACACC 5346  
QY 121 AGCACCGCACCTGGGCTTGGCCACTACAATAACCACTCTACAAGCAAAATCTCCAGT 180  
Db 5347 AGCACCGCACCTGGGCTTGGCCACTACAATAACCACTCTACAAGCAAAATCTCCAGT 5406  
QY 181 GCTTCAACCGGGGCCAGCAACCACTACTTCCGGCTACAGACCCCTCGGGGTAT 240  
Db 5407 GCTTCAACCGGGGCCAGCAACCACTACTTCCGGCTACAGACCCCTCGGGGTAT 5466  
QY 241 TTTGATTTCAACAGATTCACATGCGCATTTTACCAAGTGAATGGCAGAGCTCATCAAC 300  
Db 5467 TTTGATTTCAACAGATTCACATGCGCATTTTACCAAGTGAATGGCAGAGCTCATCAAC 5526  
QY 301 AACAAATGGGATTCGGGCCCAAGAGACTCAACTTCAAACTTTCAACATCCAAAGTCAAG 360  
Db 5527 AACAAATGGGATTCGGGCCCAAGAGACTCAACTTCAAACTTTCAACATCCAAAGTCAAG 5586  
QY 361 GAGGTCAACAGCAATGATGGCGTCAACCAATCGCTTAATAACCTTACAGCAGCGTTCAA 420  
Db 5587 GAGGTCAACAGCAATGATGGCGTCAACCAATCGCTTAATAACCTTACAGCAGCGTTCAA 5646  
QY 421 GTCTTCTCGGACTCGGAGTACAGACTTCCGATTCGATCTCGGCTCTCGGCACAGGGCTGC 480  
Db 5647 GTCTTCTCGGACTCGGAGTACAGACTTCCGATTCGATCTCGGCTCTCGGCACAGGGCTGC 5706  
QY 481 CTCCCTCCGTTCCCGGGGAGCGTGTTCATGATTCGCAATACCGCTACCTGACGCTCAAC 540  
Db 5707 CTCCCTCCGTTCCCGGGGAGCGTGTTCATGATTCGCAATACCGCTACCTGACGCTCAAC 5766



QY 541 AATGCGAGCAAGCGGTGGAGCTTCACTCTTTTACTGCTGGAATATTTCCCTTCTCAG 600  
DB |||||||  
QY 5767 AATGCGAGCAAGCGGTGGAGCTTCACTCTTTTACTGCTGGAATATTTCCCTTCTCAG 5826  
DB |||||||  
QY 601 ATGTCTGAGAACGGGCAACAACCTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTCCAC 660  
DB |||||||  
QY 5827 ATGTCTGAGAACGGGCAACAACCTTTACCTTCAGCTACACCTTTTGAGGAAGTGCCTTTCCAC 5886  
DB |||||||  
QY 661 AGCAGTACGCGCAGCAGCAGCAGCTTGACCGGCTGATGAATCCTCTCATCGACCAATAC 720  
DB |||||||  
QY 5887 AGCAGTACGCGCAGCAGCAGCAGCTTGACCGGCTGATGAATCCTCTCATCGACCAATAC 5946  
DB |||||||  
QY 721 CTGTATTACCTCAACAGAACTCAAAATCAGTCGCGAAGTGCCTTCAAAACAGAACTTGCTG 780  
DB |||||||  
QY 5947 CTGTATTACCTCAACAGAACTCAAAATCAGTCGCGAAGTGCCTTCAAAACAGAACTTGCTG 6006  
DB |||||||  
QY 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGGCCCAAAACTGGCTACTCGACCC 840  
DB |||||||  
QY 6007 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGGCCCAAAACTGGCTACTCGACCC 6066  
DB |||||||  
QY 841 TGTATTGCGCAGCAGCGCTTTCTAAACAAAACAGACAAACAAACAGCAATTTTACC 900  
DB |||||||  
QY 6067 TGTATTGCGCAGCAGCGCTTTCTAAACAAAACAGACAAACAAACAGCAATTTTACC 6126  
DB |||||||  
QY 901 TGGACTGGTGTCTCAAAATATTAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 960  
DB |||||||  
QY 6127 TGGACTGGTGTCTCAAAATATTAACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 6186  
DB |||||||  
QY 961 GCTATGGCTCTCACAAAACAGCAGAAAGCAAGATTTCTTTCCATGAGCGGTGTCATGATT 1020  
DB |||||||  
QY 6187 GCTATGGCTCTCACAAAACAGCAGAAAGCAAGATTTCTTTCCATGAGCGGTGTCATGATT 6246  
DB |||||||  
QY 1021 TTTGGAAAAGAGAGCGCCGGAGCTTCAAAACATGTCAATTTGGAACAATGTCAATTAACAGAC 1080  
DB |||||||  
QY 6247 TTTGGAAAAGAGAGCGCCGGAGCTTCAAAACATGTCAATTTGGAACAATGTCAATTAACAGAC 6306  
DB |||||||  
QY 1081 GAAGAGGAATTAAGCCCTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140  
DB |||||||  
QY 6307 GAAGAGGAATTAAGCCCTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 6366  
DB |||||||  
QY 1141 AATTTCCAGAGCAGCAGCAGACCTTGGCAGCGGAGATGTGATCTATGGGAGCATTA 1200  
DB |||||||  
QY 6367 AATTTCCAGAGCAGCAGCAGACCTTGGCAGCGGAGATGTGATCTATGGGAGCATTA 6426  
DB |||||||  
QY 1201 CTTGGCATGGTGTGGCAAGATAGAGAGCTGTACCTGCAAGGTCCCATTTTGGGCCAAAATT 1260  
DB |||||||  
QY 6427 CTTGGCATGGTGTGGCAAGATAGAGAGCTGTACCTGCAAGGTCCCATTTTGGGCCAAAATT 6486  
DB |||||||  
QY 1261 CCTCACAGATGGACACTTTCACCGCTCTCCTCTTATGGCGGCTTGGACTCAAGAAC 1320  
DB |||||||  
QY 6487 CCTCACAGATGGACACTTTCACCGCTCTCCTCTTATGGCGGCTTGGACTCAAGAAC 6546  
DB |||||||  
QY 1321 CGCGCTCTCAGATCCTCATCAAAAACAGCGCTGTTCCTGCGAATCCTCCGGCGGAGTTT 1380  
DB |||||||  
QY 6547 CGCGCTCTCAGATCCTCATCAAAAACAGCGCTGTTCCTGCGAATCCTCCGGCGGAGTTT 6606  
DB |||||||  
QY 1381 TCAGCTACAAAGTTTGCCTTCAATCATCA CCAATATCTCA CAGGA CAAAGTGAAGTGGAA 1440  
DB |||||||  
QY 6607 TCAGCTACAAAGTTTGCCTTCAATCATCA CCAATATCTCA CAGGA CAAAGTGAAGTGGAA 6666  
DB |||||||  
QY 1441 ATTGAATGGAGCTGCAGAAAGAAACAGCAGCGCTGGAAATCCCGAAGTGCAGTACACA 1500  
DB |||||||  
QY 6667 ATTGAATGGAGCTGCAGAAAGAAACAGCAGCGCTGGAAATCCCGAAGTGCAGTACACA 6726  
DB |||||||  
QY 1501 TCCAAATTATGCAAAATCTGCAACAGTTGATTTTACTGTGGACAAATGACCTTTTACT 1560  
DB |||||||  
QY 6727 TCCAAATTATGCAAAATCTGCAACAGTTGATTTTACTGTGGACAAATGACCTTTTACT 6786  
DB |||||||  
QY 1561 GAGCCTCGCCCAATTTGGCACCCGTTTACCTTTACCCGTCCTCTGTA 1604  
DB |||||||  
QY 6787 GAGCCTCGCCCAATTTGGCACCCGTTTACCTTTACCCGTCCTCTGTA 6830  
DB |||||||

RESULT 14  
CS073493  
LOCUS CS073493  
DEFINITION Sequence 81 from Patent WO2005033321.  
ACCESSION CS073493  
VERSION CS073493.1 GI:63090472  
KEYWORDS .  
SOURCE unidentified  
ORGANISM unidentified  
unclassified.  
REFERENCE 1  
AUTHORS Wilson,J.M., Gao,G., Alvita,M.R. and Vandenbergh,L.H.  
TITLE Adeno-associated virus (aav) clades, sequences, vectors containing same, and uses therefor  
JOURNAL Patent: WO 2005033321-A 81 14-APR-2005;  
The Trustees of the University of Pennsylvania (US)  
FEATURES  
Location/Qualifiers  
1..2211  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="new AAV serotype, clone hu.44"  
ORIGIN  
Query Match 99.8%; Score 1600.8; DB 6; Length 2211;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1602; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCCGCGAGTGGGT 60  
DB |||||||  
QY 607 ATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCCGCGAGTGGGT 666  
DB |||||||  
QY 61 AATGCTCAGGAATTTGGCATTTGCGATTTCCATGGCTGGGGCAGAGTCAATCACCAC 120  
DB |||||||  
QY 667 AATGCTCAGGAATTTGGCATTTGCGATTTCCATGGCTGGGGCAGAGTCAATCACCAC 726  
DB |||||||  
QY 121 AGCACCCGACCTGGGCGCTTGGCCACCTACAAATACCACTCTACAAGCAAAATCTCCAGT 180  
DB |||||||  
QY 727 AGCACCCGACCTGGGCGCTTGGCCACCTACAAATACCACTCTACAAGCAAAATCTCCAGT 786  
DB |||||||  
QY 181 GCTTCAACGGGGCGCGAGCAACGACAAACCACTACTTCGGCTACAGCACCCCTCGGGGTAT 240  
DB |||||||  
QY 787 GCTTCAACGGGGCGCGAGCAACGACAAACCACTACTTCGGCTACAGCACCCCTCGGGGTAT 846  
DB |||||||  
QY 241 TTTGATTTCAACAGATTTGCACTGCCACTTTTCAACAGTACTGGCAGGACTCATCAAC 300  
DB |||||||  
QY 847 TTTGATTTCAACAGATTTCCACTGGCACTTTTCCACAGTACTGGCAGGACTCATCAAC 906  
DB |||||||  
QY 301 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 360  
DB |||||||  
QY 907 AACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 966  
DB |||||||  
QY 361 GAGGTCAAGCAAGATGATGGCGTCAACACCATCGCTAATAACCTTACCAGCAGGTCAA 420  
DB |||||||  
QY 967 GAGGTCAAGCAAGATGATGGCGTCAACACCATCGCTAATAACCTTACCAGCAGGTCAA 1026  
DB |||||||  
QY 421 GTCTCTCGGACTCGGAGTACAGCTTCGTAAGTCTCTGGCTCTGCGCAGCAGGCTGC 480  
DB |||||||  
QY 1027 GTCTCTCGGACTCGGAGTACAGCTTCGTAAGTCTCTGGCTCTGCGCAGCAGGCTGC 1086  
DB |||||||  
QY 481 CTCCCTCGGTTCCCGCGGAGCGTGTTCATGATTCGCAATACGGCTACTGACGCTCAAC 540  
DB |||||||  
QY 1087 CTCCCTCGGTTCCCGCGGAGCGTGTTCATGATTCGCAATACGGCTACTGACGCTCAAC 1146  
DB |||||||  
QY 541 AATGGCAGCAAGCGGTGGAGCTTCACTCTTTTACTGCTTGGAAATATTTCCCTTCTCAG 600  
DB |||||||  
QY 1147 AATGGCAGCAAGCGGTGGAGCTTCACTCTTTTACTGCTTGGAAATATTTCCCTTCTCAG 1206  
DB |||||||  
QY 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCCAC 660  
DB |||||||  
QY 1207 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGGGAAGTGCCTTTCCAC 1266  
DB |||||||  
QY 661 AGCAGTACGCGCAGCAGCAGCGCTGGAGCGCTGATGAATCCTCTCATCGACCAATAC 720  
DB |||||||

Db 1267 AGCAGCTACGCGCACAGCCAGAGCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 1326  
Qy 721 CTGTATTACCTGACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAGGACTTGGCTG 780  
Db 1327 CTGTATTACCCGAACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAGGACTTGGCTG 1386  
Qy 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACTCTGGACCC 840  
Db 1387 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACTCTGGACCC 1446  
Qy 841 TGTATTCCGAGAGCGCGGTTCTTAAACAAACAGACAAACAAACAGCAATTTTACC 900  
Db 1447 TGTATTCCGAGAGCGCGGTTCTTAAACAAACAGACAAACAAACAGCAATTTTACC 1506  
Qy 901 TGGACTGGTCTTCAAAATATACCTCAATGGCGGTGAATCCATCATCAACCTCGGCAT 960  
Db 1507 TGGACTGGTCTTCAAAATATACCTCAATGGCGGTGAATCCATCATCAACCTCGGCAT 1566  
Qy 961 GCTATGGCTTCACAAAGACGACGAAGCAAGTCTTCCATGAGCGGTGTCATGATT 1020  
Db 1567 GCTATGGCTTCACAAAGACGACGAAGCAAGTCTTCCATGAGCGGTGTCATGATT 1626  
Qy 1021 TTTGAAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATGGACAATGTCATGATTACAGAC 1080  
Db 1627 TTTGAAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATGGACAATGTCATGATTACAGAC 1686  
Qy 1081 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1140  
Db 1687 GAAGAGAAATTAAGCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTGGCAGTC 1746  
Qy 1141 AATTTCCAGAGCAGCAGACAGACCTTGGACCGGAGATGTGCTGATGGAGCATTA 1200  
Db 1747 AATTTCCAGAGCAGCAGCAGACCTTGGACCGGAGATGTGCTGATGGAGCATTA 1806  
Qy 1201 CCTGGCATGGTGGCAAGTAGAGACGTGTACCTGCAGGGTCCCATTTGGGCCAAAAAT 1260  
Db 1807 CCTGGCATGGTGGCAAGTAGAGACGTGTACCTGCAGGGTCCCATTTGGGCCAAAAAT 1866  
Qy 1261 CCTCACAGATGGACACTTTTACCCGTCTCTTATGGCGGCTTTGGACTCAGAAC 1320  
Db 1867 CCTCACAGATGGACACTTTTACCCGTCTCTTATGGCGGCTTTGGACTCAGAAC 1926  
Qy 1321 CGCCCTCTCAGATCTCATCAAAAACACCCCTGTTCTTCGCAATCCTCCGCGGAGTTT 1380  
Db 1927 CGCCCTCTCAGATCTCATCAAAAACACCCCTGTTCTTCGCAATCCTCCGCGGAGTTT 1986  
Qy 1381 TCAGCTACAAAGTTTCTTCATTCATCAACCAATCTCCACAGGACAAAGTGTGGAA 1440  
Db 1987 TCAGCTACAAAGTTTCTTCATTCATCAACCAATCTCCACAGGACAAAGTGTGGAA 2046  
Qy 1441 ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCGGAAGTGCAGTACAC 1500  
Db 2047 ATTGAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCGGAAGTGCAGTACAC 2106  
Qy 1501 TCCAATTTGCAAAATCTGCAACGTTGATTTTACTGTGGACAAAGTGGACTTTTACT 1560  
Db 2107 TCCAATTTGCAAAATCTGCAACGTTGATTTTACTGTGGACAAAGTGGACTTTTACT 2166  
Qy 1561 GAGCCTCGCCCATTTGGCACCCGTTACTTTACCCGCTCCCTCTGTA 1604  
Db 2167 GAGCCTCGCCCATTTGGCACCCGTTACTTTACCCGCTCCCTCTGTA 2210

RESULT 15  
AY530607  
LOCUS AY530607 2211 bp DNA linear VRL 24-JUN-2004  
DEFINITION Adeno-associated virus isolate hu.44 capsid protein VP1 (cap) gene, complete cds.  
ACCESSION AY530607  
VERSION AY530607.1 GI:46487860  
KEYWORDS Adeno-associated virus

ORGANISM Adeno-associated virus  
REFERENCE 1 (bases 1 to 2211)  
AUTHORS Gao, G., Vandenberghe, L.H., Alvira, M.R., Lu, Y., Calcedo, R., Zhou, X. and Wilson, J.M.  
TITLE Clades of Adeno-associated viruses are widely disseminated in human tissues  
JOURNAL J. Virol. 78 (12), 6381-6388 (2004)  
PUBMED 15163731  
REFERENCE 2 (bases 1 to 2211)  
AUTHORS Gao, G., Vandenberghe, L.H., Alvira, M.R., Lu, Y., Calcedo, R., Zhou, X. and Wilson, J.M.  
TITLE Direct Submission  
JOURNAL Submitted (20-JAN-2004) Gene Therapy Program, Division of Medical Genetics, Department of Medicine, University of Pennsylvania, 3601 Spruce Street, Wistar Institute, Philadelphia, PA 19104, USA  
FEATURES  
source  
1..2211  
/organism="Adeno-associated virus"  
/mol\_type="genomic DNA"  
/isolate="hu.44"  
/isolation\_source="human tissue"  
/db\_xref="taxon:272636"  
/country="USA"  
/note="Clade A"  
1..2211  
/gene="cap"  
1..2211  
/gene="cap"  
/codon\_start=1  
/product="capsid protein VP1"  
/protein\_id="AAS99292.1"  
/translation="MAADGYLPDMLDLSGIRQWKLRPGLPPPPKPAERHKDDSRG  
LVLPQYKYLGPNGLDKGEVNEADAALHDKAYDROLDSDGNVLYKNHADAERQIG  
RUKEDTSTGGNLRGAVFQAKRVLPLGLVERGAETA PGKRPVQSGPQGPSSSSGIG  
KTGQPAKKRLNFGQTGDSVPDPLGEPATPAAVGFTTMASSGGGAMADNNSGA  
DVGASGNHWCDDSTWGLDRVITRTWALFTYNNHLYKQISSASTGASNDNHYFY  
STPMGYFDNRFHCFSPRDMQRLNNNNFRKLNFKLNIQVKEVTNDGVTITIA  
NNLTSTGVQFSDSEYQLPVLGSAHQGLPPFPADVPMIPOYGYLTNNNGSOAVGRSS  
FYCLYFYSQMLRTGNNFTESYTFEEVPHSVSAHSQSLDRMLNPLIDOLLYVPRTO  
NQSQAQNKDLFSSRGSPAGMSVQPNMLPGCYQQRQVSKTKTDNNNNFTWTGASK  
YLNNGRESIINFGTAMASHKDDKFFPMGSMVIFGKSAGASANTALDNVMTDESEI  
KATNPVATERFGTVAWFSQSDTDPDGHAMGALPGMWQGRDYLQGPWTAKTIPH  
TDGHFSPPLMGGFGLNKPPPOILLIKNTVPANPAEFSATKFPASFIQYSTQGVSE  
IEMELQKNSKRWNPVQYTSNYSANVDFVTNNGLYTEPRPIGTRYLRL"

Query Match 99.8%; Score 1600.8; DB 13; Length 2211;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1602; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGGCTTACAGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGGAGTGGGT 60  
Db 607 ATGGCTTACAGCGGTGGCGCACCAATGGCAGACAATAACGAAGGCGCGGAGTGGGT 666  
Qy 61 AATGCTCTAGGAAATTTGGCATTCCCATGCTGGCGACAGAGTCAACACC 120  
Db 667 AATGCTCTAGGAAATTTGGCATTCCCATGCTGGCGACAGAGTCAACACC 726  
Qy 121 AGCACCCGACCTGGGCGCTTGGCCACCTACAATAACACCTCTTACAAGCAATCTCCAGT 180  
Db 727 AGCACCCGACCTGGGCGCTTGGCCACCTACAATAACACCTCTTACAAGCAATCTCCAGT 786  
Qy 181 GCTTCAACGGGGCCAGCAACCGACCAACCACTACTTCGGCTACAGCACCCCTGGGGGTAT 240  
Db 787 GCTTCAACGGGGCCAGCAACCGACCAACCACTACTTCGGCTACAGCACCCCTGGGGGTAT 846  
Qy 241 TTTGATTTCAACAGATTCCACTGCCACTTTTACCACAGTGACTGGCAGGACTCATCAAC 300  
Db 847 TTTGATTTCAACAGATTCCACTGCCACTTTTACCACAGTGACTGGCAGGACTCATCAAC 906  
Qy 301 AACAAATGGGGATTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAAGTCAAG 360

Db 907 AACAAATGGGGATTTCGGGCCAAGAGACTCAACTCTTCAACATCCAGTCAAG 966  
Qy 361 GAGGTACAGCAAAATGATGGCGTCACAAACCATTCGCTAATAACCTTTACAGCAACGGTTCAA 420  
Db 967 GAGGTACAGCAAAATGATGGCGTCACAAACCATTCGCTAATAACCTTTACAGCAACGGTTCAA 1026  
Qy 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTAGTCTCTCGGCTCTGCGCACAGGCGTGC 480  
Db 1027 GTCTTCTCGGACTCGGAGTACAGCTTCGGTAGTCTCTCGGCTCTGCGCACAGGCGTGC 1086  
Qy 481 CTCCTCTCGGTTCCCGCGGAGCGTGTTCATGATTCGCGAATACGGCTACCTGACGCTCAAC 540  
Db 1087 CTCCTCTCGGTTCCCGCGGAGCGTGTTCATGATTCGCGAATACGGCTACCTGACGCTCAAC 1146  
Qy 541 AATGCGAGCAAGCGGTGGAGCGTTTCATCTTTTACTGCTTGGAAATATTTCCCTTCTCAG 600  
Db 1147 AATGCGAGCAAGCGGTGGAGCGTTTCATCTTTTACTGCTTGGAAATATTTCCCTTCTCAG 1206  
Qy 601 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCAC 660  
Db 1207 ATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCTTTCAC 1266  
Qy 661 AGCAGCTACCGCACAGCGCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720  
Db 1267 AGCAGCTACCGCACAGCGCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 1326  
Qy 721 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAAAGGACTTGCTG 780  
Db 1327 CTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAAAGGACTTGCTG 1386  
Qy 781 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACCTGGACCC 840  
Db 1387 TTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACCTGGACCC 1446  
Qy 841 TGTATTCCGCGACAGCGCGTTCCTAAACAAAAACAGACAAACAAACAGCAATTTTACC 900  
Db 1447 TGTATTCCGCGACAGCGCGTTCCTAAACAAAAACAGACAAACAAACAGCAATTTTACC 1506  
Qy 901 TGGACTGGTGTCTTCAAAATATACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 960  
Db 1507 TGGACTGGTGTCTTCAAAATATACCTCAATGGCGGTGAATCCATCATCAACCTGGCACT 1566  
Qy 961 GCTATGGCCCTCACAAAGACGACGAGACAGATGCTTTCCCATGAGCGGTGTATGATT 1020  
Db 1567 GCTATGGCCCTCACAAAGACGACGAGACAGATGCTTTCCCATGAGCGGTGTATGATT 1626  
Qy 1021 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGCAATGTCTATGATTAACAGAC 1080  
Db 1627 TTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGCAATGTCTATGATTAACAGAC 1686  
Qy 1081 GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140  
Db 1687 GAAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1746  
Qy 1141 AATTTCCAGACGACGACAGACCCCTGCGA CCGGAGATGTGCATGCTATGGGAGCATTA 1200  
Db 1747 AATTTCCAGACGACGACAGACCCCTGCGA CCGGAGATGTGCATGCTATGGGAGCATTA 1806  
Qy 1201 CCTGGCATGTGTGGCAAGATAGAGAGGTGATCCTGCGAGGTCCCATTTGGGCCAAAATT 1260  
Db 1807 CCTGGCATGTGTGGCAAGATAGAGAGGTGATCCTGCGAGGTCCCATTTGGGCCAAAATT 1866  
Qy 1261 CCTCACAGATGGACACTTTTCAACCGGTCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 1320  
Db 1867 CCTCACAGATGGACACTTTTCAACCGGTCTCTCTTATGGGCGGCTTTGGACTCAAGAAC 1926  
Qy 1321 CGGCTCTCTCAGATCTCATCAAAAACAGCCCTGTTCTTCCGGAATCCTCGGCGGAGTTT 1380  
Db 1927 CGGCTCTCTCAGATCTCATCAAAAACAGCCCTGTTCTTCCGGAATCCTCGGCGGAGTTT 1986  
Qy 1381 TCAGCTACAAAGTTTCCTTCACTTCACTCA CCAATACCTCCACAGGACAAGTGTGGAA 1440

Db 1987 TCAGCTACAAAGTTTGTCTTCATTTCATCCCAATACTCCACAGGACAAAGTGTGTGGAA 2046  
Qy 1441 ATTGAATGGGAGCTGCGAGAAAGAAAAACAGCAAGCGCTGGAAATCCCGAAAGTGCAGTACACA 1500  
Db 2047 ATTGAATGGGAGCTGCGAGAAAGAAAAACAGCAAGCGCTGGAAATCCCGAAAGTGCAGTACACA 2106  
Qy 1501 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTTATCT 1560  
Db 2107 TCCAAATTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACAATGGACTTTTATCT 2166  
Qy 1561 GAGCCTCGCCCAATTTGGCACCCGTTACCTTACCCGTCCCTGTGA 1604  
Db 2167 GAGCCTCGCCCAATTTGGCACCCGTTACCTTACCCGTCCCTGTGA 2210

Search completed: November 29, 2005, 00:05:36

Job time : 5865.51 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 18:24:29 ; Search time 7013.94 Seconds  
(without alignments)  
14735.325 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_2223\_4431  
Perfect score: 2209  
Sequence: 1 atggctgcgatgttatct.....ttacctaccgtccctcgt 2209

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: gb\_est1:\*
- 2: gb\_est2:\*
- 3: gb\_est3:\*
- 4: gb\_est4:\*
- 5: gb\_est5:\*
- 6: gb\_est6:\*
- 7: gb\_est7:\*
- 8: gb\_est8:\*
- 9: gb\_est9:\*
- 10: gb\_est10:\*
- 11: gb\_est11:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	180.8	8.2	264	7	CO892248 BovGen_20
C 2	125.8	5.7	195	7	CO888893 BovGen_17
C 3	76.2	3.4	753	9	BH115587 RPI-24-3
C 4	50.6	2.3	803	9	BZ265284 CH230-375
C 5	48	2.2	367	3	BJ432209 BJ432209
C 6	47	2.1	581	5	BW335572 BW335572
C 7	45.6	2.1	681	10	AL133930 Tetraodon
C 8	45.6	2.1	869	7	CK159167 FGAS04056
C 9	45.4	2.1	581	6	CF569196 EST057 Su
C 10	45.4	2.1	581	11	CN304QNN
C 11	45.2	2.0	1075	10	CW931784
C 12	45	2.0	289	1	AI796567
C 13	45	2.0	322	1	AW206887 UI-H-Bil-
C 14	45	2.0	356	2	BG170384 602322706
C 15	45	2.0	451	1	AI439550 tc90C03.x
C 16	45	2.0	454	1	AI240019 gh33h11.x
C 17	45	2.0	472	1	AA806398 oc27a06.s
C 18	45	2.0	479	1	AI566324 tq71b07.x
C 19	45	2.0	486	3	BM352611 ig68h10.x
C 20	45	2.0	488	1	AW181950 xj69a04.x
C 21	45	2.0	498	2	BF194970 7091b09.x
C 22	45	2.0	504	1	AI989403 wt80g09.x

C 23	45	2.0	511	1	AI018231
C 24	45	2.0	535	2	BF732474
C 25	45	2.0	537	1	AI761007
C 26	45	2.0	551	1	AA421597
C 27	45	2.0	612	1	AI523558
C 28	45	2.0	620	5	BX349872
C 29	45	2.0	664	2	BF058781
C 30	45	2.0	665	2	BE903379
C 31	45	2.0	762	1	AI373205
C 32	45	2.0	910	10	CNS0060N
C 33	45	2.0	1362	4	BC020427
C 34	44.8	2.0	199	1	AA864674
C 35	44.8	2.0	556	3	BI948203
C 36	44.8	2.0	821	3	BI948601
C 37	44.8	2.0	830	8	CV764236
C 38	44.8	2.0	1101	8	DR736539
C 39	44	2.0	362	1	AI870396
C 40	44	2.0	381	1	AI240354
C 41	44	2.0	579	1	AI762877
C 42	44	2.0	622	6	CB870960
C 43	44	2.0	925	10	CNS0091P
C 44	43.8	2.0	512	5	CA024678
C 45	43.8	2.0	596	1	AV601091

ALIGNMENTS

RESULT 1  
CO892248/c  
LOCUS  
DEFINITION BovGen 20573 normal cattle brain Bos taurus cdna clone  
RZPDp1056M0360Q 5', mRNA sequence.  
ACCESSION CO892248  
VERSION CO892248.1 GI:51922548  
KEYWORDS EST.  
SOURCE Bos taurus (cow)  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;  
Pecora; Bovidae; Bovinae; Bos.  
REFERENCE  
Hennig S., Janitz M., Herwig R. and Williams J.  
Generation, annotation, evolutionary analysis and database  
integration of 14969 cattle EST clusters  
Unpublished (2004)  
JOURNAL  
COMMENT  
Contact: Hennig S  
Laboratory 123, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Innestr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1612  
Fax: +49 30 8413 1380  
Email: hennigmolgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting  
(ONFP) to reduce sequencing redundancy. According to the ONFP  
procedure, clones that display the same hybridisation matrix with a  
battery of 200 8mer oligonucleotides are grouped into clusters. One  
clone per ONFP cluster was selected for sequencing. cDNA clones and  
filters are distributed via Deutsches Ressourcenzentrum fuer  
Genomforschung GmbH (<http://www.rzpd.de>).  
PCR Primers  
FORWARD: 5' CCCAGGCTTTACACTTTATGCTTCGGGTCG 3' (M13RSP) 5'-seq  
BACKWARD: 5' GCTATTACCCAGCTGGCAAGGGGATGTG 3' (M13FSP) 3'-seq  
Seq primer: 5'-CCGCTCCGGATCCCGGT-3' (M13RSP).

FEATURES  
source

Location/Qualifiers  
1..264  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/clone="RZPDp1056M0360Q"  
/sex="female"  
/tissue\_type="brain tissue"  
/dev\_stage="adult brain"

```

/clone_lib="normal cattle brain"
/site="Organ: brain; Vector: pSport1; Site 1: NotI;
Site 2: SalI; Random primed and directionally cloned in
pSport1 vector using NotI
(5'-pGACTAGTCTAGATCGGAGCGCGCCG (7)15-3' and SalI 5'-
TCGACCCACGGGTCCG-3' adapters (Gibco BRL))"

ORIGIN
Query Match      8.2%; Score 180.8; DB 7; Length 264;
Best Local Similarity 80.3%; Pred. No. 1.2e-39;
Matches 212; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1847 GTCCCATTTGGCCAAAATTCCTCACACAGATGGACATTTCCACCGTCTCTCTTATGG 1906
      |||||
Db 264 GTCCCATTTGGCAATGATTCACACTCGGACGGAATTTTCCACCCCTCTCCCTTCATGG 205
      |||||
QY 1907 GCGGCTTTGGACTCAAGAACCCGCTCTCTCAGATCCTCATCAAAACACGCGTGTTCCTG 1966
      |||||
Db 204 GGGGATTTGGACTTAATCACCTCTCTCCACAGATTTTCATCAAGAACGCGGTTCTCTG 145
      |||||
QY 1967 CQAATCTCTCGCGGAGTTTCAGCTACAAAGTTTGCTTCATTCATCACCAGTACTCCA 2026
      |||||
Db 144 CGAGTCTCTCGAGCAGCTTCAGTGGCGCAAGTTTGCTTCATTCATCAGACAGTACTCCA 85
      |||||
QY 2027 CAGGACAAAGTGTGGAATTTGAATGGAGCTCGAGAAAGAAACACAGCAAGCGCTGGA 2086
      |||||
Db 84 CGGTCAGTGTAGCTGGAGATTGATGGAGCTGGGAGGAAACACAGCAAGCGCTGGA 25
      |||||
QY 2087 ATCCCGAAGTGCAGTACATCCA 2110
      |||||
Db 24 ATCCCGAATTCAGTACACTTCCA 1
      |||||

```

```

RESULT 2
CO888893/c
LOCUS      195 bp      mRNA      linear      EST 01-SEP-2004
DEFINITION BovGen.17218 normal cattle brain Bos taurus cDNA clone
ACCESSION  RZPD0105600960Q 5', mRNA sequence.
VERSION     CO888893.1 GI:51819178
KEYWORDS   EST.
SOURCE     Bos taurus (cow)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE  Hennig,S., Janitz,M., Herwig,R. and Williams,J.
            Generation, annotation, evolutionary analysis and database
            integration of 14969 cattle EST clusters
JOURNAL    Unpublished (2004)
AUTHORS    Contact: Hennig S
            Laboratory 123, dept.Lehrach
            Max-Planck-Institut fuer Molekulare Genetik
            Ihnestr.63-73, D-14195 Berlin, Germany
            Tel: +49 30 8413 1612
            Fax: +49 30 8413 1380
            Email: hennig@molgen.mpg.de
COMMENT    The library was characterised by oligonucleotide fingerprinting
            (ONFP) to reduce sequencing redundancy. According to the ONFP
            procedure, clones that display the same hybridisation matrix with a
            battery of 200 8mer oligonucleotides are grouped into clusters. One
            clone per ONFP cluster was selected for sequencing. cDNA clones and
            filters are distributed via Deutsches Ressourcenzentrum fuer
            Genomforschung GmbH (http://www.rzpd.de).
PCR Primers
FORWARD: 5' CCCGAGCTTTACACTTTATGCTTCGGCTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GCTATTACCGAGCTGGCGAAGGGGATGG 3' (M13FSP) 3'-seq
Seq primer: 5'-CCGGTCGGAATTCCTCCGGT-3' (M13RSP).

Location/Qualifiers
1. .195
/organism="Bos taurus"
/mol_type="mRNA"

FEATURES
source

```

```

/db_xref="taxon:9913"
/clone="RZPD0105600960Q"
/sex="female"
/tissue_type="brain tissue"
/dev_stage="adult brain"
/clone_lib="normal cattle brain"
/site="Organ: brain; Vector: pSport1; Site 1: NotI;
Site 2: SalI; Random primed and directionally cloned in
pSport1 vector using NotI
(5'-pGACTAGTCTAGATCGGAGCGCGCCG (7)15-3' and SalI 5'-
TCGACCCACGGGTCCG-3' adapters (Gibco BRL))"

ORIGIN
Query Match      5.7%; Score 125.8; DB 7; Length 195;
Best Local Similarity 78.2%; Pred. No. 4.9e-24;
Matches 151; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 2017 CAATACTCCACAGACAAAGTGTGGAATTTGAATGGAGCTGCAGAAAGAAAACAGC 2076
      |||||
Db 194 CAGTACTCCACGGGACAGGTGAGGTGGAGATCGATGGGAGTTGCAGAGGTAAACAGC 135
      |||||
QY 2077 AAGCGCTGGAATCCGAAAGTGCAGTACACATCCCAATTAATGCAAAATTCGCAACGTTGAT 2136
      |||||
Db 134 AAACGGTGGAAATCCGAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATGTGAT 75
      |||||
QY 2137 TTTACTGTGGACAAATGGAGCTTTATACTGAGCTCGCCCATTTGGCACCCTTACCTT 2196
      |||||
Db 74 TTTAATGTGGACATTAATGCGGTGTATTTCAGAGCTCGCCCATTTGGCACCAGATACCTG 15
      |||||
QY 2197 ACCCGTCCCTGT 2209
      |||||
Db 14 ACTCGTAATGTGT 2

RESULT 3
BH115587
LOCUS      753 bp      DNA      linear      GSS 19-JUL-2001
DEFINITION RPCI-24-358F16.TV RPCI-24 Mus musculus genomic clone
ACCESSION  RPCI-24-358F16, genomic survey sequence.
VERSION     BH115587
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsagaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
            1 (bases 1 to 753)
            Mouse BAC End Sequences from Library RPCI-24
            Other GSSs: RPCI-24-358F16.TJ
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pjejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Place: 358 row: F column: 16
            Seq primer: T7
            Class: BAC ends.
            Location/Qualifiers
            1. .753
            /organism="Mus musculus"
            /mol_type="genomic DNA"

FEATURES
source

```

```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-350F16"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/notes="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI; The
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

ORIGIN
Query Match      3.4%; Score 76.2; DB 9; Length 753;
Best Local Similarity 51.6%; Pred. No. 1e-09;
Matches 268; Conservative 0; Mismatches 223; Indels 28; Gaps 3;

QY 266 CGTACCTCGGTATATACCGCCGAGCGGAGTTTTCAGGAGCGCTGCGAAGAGATACGT 325
DB 207 CTTGTCTCCATAAATACACAGAGCAGAGTTTCAGGAGAAACTCCAAAGTACATAC 266
QY 326 CTTTGGGGCACCTCGGGCGGAGCAGTCTTCAGGCGCAAGAGCGGTTCTCGAACCTC 385
DB 267 TTTTTTTGGCA-----GGAACCGTGCCAATGCCAAAGAAAGGGTTCTCAAACCCA 318
QY 386 TCGGCTCTGTTGAGGAGCGCTTAAGACGGCTCTCGAAAGAAACGTCCTCGGTAGAGCAGT 445
DB 319 TCGGCTTGGTCTAAGAGAGATTACGATGCTCTCTGGAGAAAT----- 363
QY 446 CGCCACAAGAGCCAGACTCTCTCTCGGGCATCGGCAAGACAGCGCAGCAGCCCGCTAAAA 505
DB 364 ----AGATTTCGACTTAAACTCTCTCCCAACAGAGAGATACACTCGCTCTTTTCAGA 419
QY 506 AGAGACTCATTTTGGTCAGACTGGCGACTCAGAGTCAGTCCCGATCCCAACCTCTCG 565
DB 420 GTTATGCCAAAAACAGTAGAGTGGAGAGCTGCTCATTTAGTCTATGACACACAGATGTA 479
QY 566 GAGAACCTCCAGCAACCCCGCTGCTGGGACCTCTACAATGGCTTTCAGGGGGTGGCG 625
DB 480 GCAGTTTCCAGCAGATATCTCCATTTAGACACTTCTATAATCTCTGGAGCTGGAGTTC 539
QY 626 CACCAATGGCAGACAATAAGAGAGCGCCGAGGAGTGGGTATGCTCTCAGGAAATTTGGC 685
DB 540 TACTAATGGCAACAACAAC- CAGACACTGATGGAGTGAGCAATGCCATGGATGATGGC 598
QY 686 ATTGCGAATTCACATGCTGGCGACAGAGTCATCACACAGCAGCAGCCGACCTGGCGCT 745
DB 599 ATTGCCAATTCGAAGTGTATGGTAGACTGAGTCAATTACCTAATCTCCCCAGACCTGGGTCC 658
QY 746 TGCCACCTACATAAACCACTCTACAAAGCAAAATCTCCA 784
DB 659 TGCCAACTCAACTAAACATTTCCACAACTATATGAACA 697

RESULT 4
BZ265284
LOCUS CH230-375A22.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
DEFINITION CH230-375A22, genomic survey sequence.
ACCESSION BZ265284
VERSION BZ265284.1 GI:23978528
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 803)
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 MboI segment
JOURNAL Unpublished (1999)

Other_GSSs: CH230-375A22.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering_information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 375 row: A column: 22
Seq primer: SP6
Class: BAC ends.

FEATURES
source
1..803
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-375A22"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"

ORIGIN
Query Match      2.3%; Score 50.6; DB 9; Length 803;
Best Local Similarity 60.6%; Pred. No. 0.02;
Matches 83; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 276 GTATAACACCGCCGAGCGAGTTTCAGGAGCGTCTGCAAGAGATACGTCTTTTGGGG 335
DB 304 GTACAGCCCTGTGCAAGCAGAAATTCAGAGAAATCCAAACAGATGCTTTTCCAGAGG 363
QY 336 CACCTCGGGGAGCAGTCTCCAGGCGGAGCGGTTCTCGAACCTCTCGGTCTGGT 395
DB 364 CAATCTCAACCAAGTACTCTTCAGGCTAAAGAGACTTCTGATCCCTTTGACCTGGT 423
QY 396 TGAGGAAGCGCTAAGA 412
DB 424 TAAAAAGTTGTCCAGA 440

RESULT 5
BJ432209/c
LOCUS BJ432209 Dictyostelium discoideum cDNA library, VF Dictyostelium
DEFINITION BJ432209
ACCESSION BJ432209
VERSION BJ432209.1 GI:19406931
KEYWORDS EST.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
REFERENCE 1 (bases 1 to 367)
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoideum at the vegetative
stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES

```



```

source
1. .367
/organism="Dictyostellium discoideum"
/mol_type="mrna"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddv18d02"
/sex="mat A"
/dev_stage="Growth phase"
/clone_lib="Dictyostellium discoideum cDNA library, VF"

ORIGIN
Query Match 2.2%; Score 48; DB 3; Length 367;
Best Local Similarity 49.6%; Pred. No. 0.085;
Matches 120; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 716 TCATCACCACGACCCGACCTGGGCTTGGCCACCTTACCAATTAACCACTTCTACAAGC 775
DB 328 TCAACAACATCACAACATCAANTACCTCTAAACCAACAACCTAGATCAACATCAACT 269
QY 776 AAATCTCAGTCTTCAACGGGGCCGACGACGACCAACCACTACTTCCGCTCAGCACCC 835
DB 268 ACATCAACTACCTCTAAACCAACAACCAACATCAACTTACATCAACTCACTACATCA 209
QY 836 CTGGGGGTATTTGATTTTCAACAGATTCCACTTCCACTTTTCCACCACTGACTGGCAGC 895
DB 208 ACTACTTCCAAACCACTACTATCAACCAACATCAACTTCTCCAAACCACTACTGCT 149
QY 896 GACTCATCAACAACAATTTGGGATTCGGCCCAAGAGACTCAACTTTCAAACTTTCAACA 955
DB 148 TCATCAACAACAACAACCTGTTTCATCAACAACAACCTGTTTCATCAACAACAACCATCA 89
QY 956 TC 957
DB 88 CC 87

RESULT 6
BM335572 581 bp mrna linear EST 27-MAY-2004
LOCUS BM335572 Yutaka Satou unpublished cDNA library, adult digestive
DEFINITION gland Clona intestinalis cDNA clone cidg851f15', mRNA sequence.
ACCESSION BM335572.1 GI:47747373
VERSION 1 (bases 1 to 581)
KEYWORDS Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
SOURCE Unpublished (2004)
ORGANISM Contact: Yutaka Satou
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4095
Fax: 81-75-705-1113
Email: yutaka@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
1. 581
/organism="Clona intestinalis"
/mol_type="mrna"
/db_xref="taxon:7719"
/clone="cidg851f15"
/tissue_type="digestive gland"
/dev_stage="adult"
/clone_lib="Yutaka Satou unpublished cDNA library, adult
digestive gland"

ORIGIN
Query Match 2.1%; Score 47; DB 5; Length 581;
Best Local Similarity 45.9%; Pred. No. 0.19;

Matches 161; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 699 GCGATTTCACATGGCTGGGCGACAGAGTCATCACACCAGCACCGACCTGGGCTTGGC 748
DB 148 GCAACTCAGAAAGCACCACCAACATCATCAACAACAACAGCAGCGGCAACAACAGCAACA 207
QY 749 CCACCTACAATAACCACTCTTACAGCAAAATCTCCAGTGCTTCAACGGGGGCCAGCAACG 808
DB 208 CCAACAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 267
QY 809 ACAACCACTACTTGGCTTACAGCACCCCTGGGGTATTTTGAATTTCAACAGATTCACCT 868
DB 268 ACAACGAGTCATCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 327
QY 869 GCCACTTTTACCAGTGACTGGCAGCGACTCATCAACAACAATAATTTGGGATTCGGGCCCA 928
DB 328 GGAACAACAACGGCACCACCAACAACAACAACAACAACAACAACAACAACAACAACAACA 387
QY 929 AGAGACTCAACTTCAAACTCTTCAACATCCCAAGTCAAGGAGTCAACGAGATGATGGCG 988
DB 388 ACAACAGAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 447
QY 989 TCACAACCATCGCTAATAACCTTACAGCAGCGTTTCAAGTCTTCTCGGACT 1039
DB 448 ACAACAACAACGCTTATCAAGAACAACAACAACAACAACAACAACAACAACAACAACA 498

RESULT 7
CNS02EOD 681 bp DNA linear GSS 01-SEP-2000
LOCUS CNS02EOD/4
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
262H14 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL193990
VERSION AL193990.1 GI:7832096
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
REFERENCE 1 Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
AUTHORS Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
PUBMED 10835645
REFERENCE 2 Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
AUTHORS Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
PUBMED 10899143
REFERENCE 3 (bases 1 to 681)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
1. .681
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
FEATURES
source

```



```
/clone="262H14"  
/clone_lib="G"  
/notes="Genoscope sequence ID : COAG262DD07LP1  
end : T7"
```

## ORIGIN

```
Query Match 2.1%; Score 45.6; DB 10; Length 681;  
Best Local Similarity 43.4%; Pred. No. 0.51;  
Matches 144; Conservative 10; Mismatches 178; Indels 0; Gaps 0;  
  
QY 637 GACAATTAACGAAGCGCGCGAGTGGGTAAATCCCTCAGGAATTTGGCATTCGATTCC 696  
DB 351 GACAATTAACGAAGCGCGAGTGGGTAAATCCCTCAGGAATTTGGCATTCGATTCC 292  
  
QY 697 ACATGGCTGGCGGACAGAGTCAATCACCACGACCGCGACCTTGGGCTTGGCCACCTAC 756  
DB 291 RACGGGACGACGACGAGCGAGGACCAACAAAACGACRACGCGCGACGACGACAC 232  
  
QY 757 AATAACCACTTCAAGCAAAATCTCCAGTGTCTTCAACGGGGCGCGACGACGACACAC 816  
DB 231 AACAACAACAACAACAAAACAAACGCAACGCAACGCGNACGCAACAAACNACACAC 172  
  
QY 817 TACTTGGCTAGACACCCCTGGGGTATTTTGAATTTCAACAGATTCCTACTGCCATTT 876  
DB 171 AAAAAACRACAAACACRACGCGACAAACAACRACAAACAAACAAACAAACAAAC 112  
  
QY 877 TCACCACTGCTGGCGAGCGACTCATCAACAACAATTTGGGGATTTCGGGCGCAAGAGACTC 936  
DB 111 ACAAACAACRACRACRACRACRACRACRACRACRACRACRACRACRACRACRAC 52  
  
QY 937 AACTTCAAACTTCTCAACATCCAGTCAAGGA 968  
DB 51 AASMACAAACRACRACRACRACRACGACGACGACGACGACGACGACGACGACGAC 20
```

## RESULT 8

```
CK159167/c  
LOCUS CK159167 869 bp mRNA linear EST 05-DEC-2003  
DEFINITION FGAS040564 Triticum aestivum FGAS: TaLts5 Triticum aestivum cDNA,  
mRNA sequence.
```

## ACCESSION

```
CK159167
```

## VERSION

```
CK159167.1 GI:38989053
```

## KEYWORDS

```
EST.
```

## SOURCE

```
Triticum aestivum (bread wheat)
```

## ORGANISM

```
Triticum aestivum
```

```
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.
```

```
1 (bases 1 to 869)
```

```
Allard,F., Crosby,W.L., Danyluk,J., Eudes,F., Frick,M., Gaudet,D.,  
Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A.,  
Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilson,D.,  
Penniket,C., Roach,J.L. and Sarhan,F.
```

```
Functional Genomics of Abiotic Stress In Wheat and Canola Crops  
Unpublished (2003)
```

```
Contact: Wm L Crosby
```

```
Bioinformatics
```

```
University of Saskatchewan, Department of Computer Science
```

```
1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
Saskatchewan, S7N 5A9, Canada
```

```
Tel: 306 966 1769
```

```
Fax: 306 966 2033
```

```
Email: fgas_este@cs.usask.ca
```

This sequence is the direct result of the Base calling software  
Phred (default parameters). It is the raw base calls. To aid in the  
identification of the high quality insert the software Lucy  
(default parameters) has been run on this sequence. Lucy identified  
the region [128,636].  
Plate: TaLts537 row: N column: 23.

## FEATURES

## source

```
1. .869  
/organisms="Triticum aestivum"  
/mol_type="mRNA"
```

```
/cultivar="Wheat line PI 178383"
```

```
/db_xref="taxon:4565"
```

```
/clone_lib="Triticum aestivum FGAS: TaLts5"
```

/note="Organ: Crown; Vector: pGEM-T; SSH (suppression  
subtractive hybridization) cDNA library from genotype  
P1178383 cold hardened at 2 C for 21 days and 49 days  
(equal amount of cDNA pooled together before subtraction,  
tester) and subtracted against genotype Norstar cold  
hardened at 2 C for 1 day (24 H) (driver). Modified Smart  
cDNA (Clontech) priming and non-directional cloning"

## ORIGIN

```
Query Match 2.1%; Score 45.6; DB 7; Length 869;  
Best Local Similarity 48.1%; Pred. No. 0.56;  
Matches 129; Conservative 0; Mismatches 139; Indels 0; Gaps 0;  
  
QY 751 ACTTACAATTAACCACTCTTACAGCAAAATCTCCAGTGTCTTCAACGGGGCGCGACGAC 810  
DB 785 AACAACCAACCAACCAACCAACGACGACGACGACGACGACGACGACGACGACGAC 726  
  
QY 811 AACCACTACTTGGCTTACAGCACCCCTGGGGTATTTTGAATTTCAACAGATTCCTACTGC 870  
DB 725 AACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 666  
  
QY 871 CACTTTTCAACCACTGTCGCGAGCTCATCAACAACAATTTGGGGATTTCGGGCGCAAG 930  
DB 665 CACCACAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 606  
  
QY 931 AGACTCACTTCAACTCTTCAACATCCAGTCAAGGAGGTACGAGGATGATGGCGTC 990  
DB 605 AACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 546  
  
QY 991 ACAACCATGCTTAATTAACCTTACCGCA 1018  
DB 545 AACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 518
```

## RESULT 9

## LOCUS

```
CF569196
```

## DEFINITION

```
CF569196 581 bp mRNA linear EST 08-SEP-2004  
EST057 Subtracted, Clontech (cat. # K1804-1) Triticum aestivum cDNA  
clone FDC57 5', mRNA sequence.
```

## ACCESSION

```
CF569196
```

## VERSION

```
CF569196.1 GI:51921536
```

## KEYWORDS

```
EST.
```

## SOURCE

```
Triticum aestivum (bread wheat)
```

## ORGANISM

```
Triticum aestivum
```

```
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.
```

```
1 (bases 1 to 581)
```

```
Xiao,X., Bai,G.H. and Carver,B.F.
```

```
Nylon Filter Arrays Reveal Differential Expression of Expressed  
Sequence Tags in Wheat Roots Under Aluminum Stress
```

```
J. Integr. Plant Biol. 47 (7), 839-848 (2005)
```

```
Contact: Guihua Bai
```

```
USDA/ARS and Department of Agronomy
```

```
Kansas State University
```

```
Manhattan, KS 66506, USA
```

```
Email: gbai@bear.agron.ksu.edu
```

```
Seq primer: M13 Forward
```

```
High quality sequence stop: 581.
```

```
Location/Qualifiers
```

```
1. .581
```

```
/organism="Triticum aestivum"
```

```
/mol_type="mRNA"
```

```
/cultivar="OK91G106"
```

```
/db_xref="taxon:4565"
```

```
/clone="FDC57"
```

```
/tissue_type="root"
```

```
/clone_lib="Subtracted, Clontech (cat. # K1804-1)"
```

```
/notes="EST from wheat (Triticum aestivum, cv. OK91G106)"
```

root in response to aluminum stress"

```
ORIGIN
Query Match          2.1%; Score 45.4; DB 6; Length 581;
Best Local Similarity 46.6%; Pred. No. 0.55;
Matches 145; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 708 CGACGAGTCATCACCACGACCGACCTGGGCTTGGCCCTACCAATACCACT 767
Db 152 CAACAACAACAACACCAACAGCAGCAGCAACAGCAACAACAGCAACAACA 211

QY 768 CTACAGCAAAATCTCCAGTGTCTTCAACGGGGGCGAGCAACCACTACTTGGCTA 827
Db 212 CAACAACAACAACAGCAACAACAACAGCAACAACAACAACAACAACAACA 271

QY 828 CAGACCCCTGGGGGTATTTTGAATTTCAAGATTCCTACCTCTTTTACACGTA 887
Db 272 CAGCAGCAACAACAGCAGCAACAACAGCAGCAGCAGCAACAGTAAACAACA 331

QY 888 CTGGCAGCACTCATCAACAACAATTTGGGGATTCGGCCCAAGAGACTCAACTTCAACT 947
Db 332 CAACAGCAGCAACAACAACAACAACAACAACAACAACAACAACAACAACA 391

QY 948 CTTCAACATCTCAAGTCAAGAGGTTCAGCAGATGATGCGGTCAACAACCATCGCTAATAA 1007
Db 392 CAGCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 451

QY 1008 CTTTACCAGCA 1018
Db 452 CAGCAACAGCA 462

RESULT 10
CN504QNN/c
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence T7 end of clone
129H24 of library G from Tetraodon nigroviridis, genomic survey
sequence.
AL302828
AL302828.1 GI:8183374
GSS; genome survey sequence.
SOURCE
Tetraodon nigroviridis
ORGANISM
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE
1 Roest Crolius H., Jaillon O., Dasilva, C., Bouneau L., Fisher, C.,
Bernot, A., Pizanes, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
10835645
2
Roest Crolius H., Jaillon O., Dasilva, C., Ozouf-Costaz, C.,
Pizanes, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
Saurin, W., Bernot, A. and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
10893143
3 (bases 1 to 983)
Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 131 91066 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.

FEATURES
source
Location/Qualifiers
1..983
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="129H24"
/clone_lib="G"
/note="Genoscope sequence ID : COBG129DD12LP1
end : 77"

ORIGIN
Query Match          2.1%; Score 45.4; DB 11; Length 983;
Best Local Similarity 47.1%; Pred. No. 0.66;
Matches 139; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 29 GCGTCGAGGACAACCTCTCTGAGGCAATTCGCGAGTGTGGGACTTGAAACCTGAGGCC 88
Db 381 GCCAGGAGCAGGACCAGGGCCAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAG 322

QY 89 CGAAGCCAAAGCCAAACAGCAAAAGCAGGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 148
Db 321 GCCAGGAGCAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAG 262

QY 149 ACAAGTACTCTCGACCTTCAACGAGCTCGACAAGAGGAGGAGGAGGAGGAGGAGGAGG 208
Db 261 GCCAGGAGCAGGACCAAGGACCAAGGACCAAGGACCAAGGAGGAGGAGGAGGAGGAGG 202

QY 209 CAGCGGCTCTGAGCAGCAAGGCTTACGACCAAGGAGCTCAAAAGCGGGTGCATTCCT 268
Db 201 ACCAGGAGCAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAG 142

QY 269 ACCTCGGATTAACCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 323
Db 141 GCCAGGAGCAGGACCAAGGACCAAGGACCAAGGACCAAGGAGGAGGAGGAGGAGGAGGAG 87

RESULT 11
CN504QNN/c
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence T7 end of clone
129H24 of library G from Tetraodon nigroviridis, genomic survey
sequence.
AL302828
AL302828.1 GI:60247281
GSS.
SOURCE
Tetraodon nigroviridis
ORGANISM
Tetraodon nigroviridis
Eukaryota; Acanthamoebidae; Acanthamoeba.
REFERENCE
1 (bases 1 to 1075)
Anderson, I.J. and Loftus, B.J.
Gene discovery in the Acanthamoeba castellanii genome
Unpublished (2004)
Contact: Iain Anderson
The Institute for Genomic Research (TIGR; www.tigr.org)
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-795-7949
Fax: 301-838-0208
Class: shotgun.
Location/Qualifiers
1..1075
/organism="Acanthamoeba castellanii"
/mol_type="genomic DNA"
/strain="Neif"
/db_xref="taxon:5755"
/clone="EDCC981"
/clone_lib="A. castellanii, 6-8 kb library from total
genomic DNA"
/note="Vector: pHOS2"

FEATURES
source
Location/Qualifiers
1..1075
/organism="Acanthamoeba castellanii"
/mol_type="genomic DNA"
/strain="Neif"
/db_xref="taxon:5755"
/clone="EDCC981"
/clone_lib="A. castellanii, 6-8 kb library from total
genomic DNA"
/note="Vector: pHOS2"

ORIGIN
Query Match          2.0%; Score 45.2; DB 10; Length 1075;
Best Local Similarity 54.1%; Pred. No. 0.78;
Matches 92; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
```

```

QY 100 GCCAACCAAGAAAGCAGACGACGGCCGGGGTCTGTGCTTCTGCTTCAAGTACCTC 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 698 GCGACGAGGACTCGAGCTGGCCGCGCAGGTGAGCGCGCTGCAGCGCGCAAGGACGG 639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 160 GGACCTTTCAACGAGACTCGACAAAGGGGGAGCCCGTCAACGGCGCGACGACGGCCCTC 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 638 CTGGCGCTCGAGTGGCCGACAGGACGAGCGCGTCAAGGAGCTCGAGAAAGCGGCCGAC 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 220 GAGCAGCAAGCGCTTACGACGACGAGCTCAAGCGGGTGACAATCCGTA 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 578 GACCTGGTCAAGGATACGAGAGGAGAAAGCAGCGCGCGCAGACCTGGA 529
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 12
A1796567/c
LOCUS      289 bp      mRNA      linear      EST 20-DEC-1999
DEFINITION wh46h04.x1 NCI_CGAP Kid11 Homo sapiens cDNA clone IMAGE:238381 3'
            similar to TR:000613 000613 T3 RECEPTOR-ASSOCIATING COFACTOR-1. [1]
            ; mRNA sequence.

```

```

ACCESSION A1796567
VERSION   A1796567.1 GI:5362030
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 289)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
           Emmert-Buck, M.D., Ph.D.

```

```

CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 326 Std Error: 0.00
Seq primer: -40UP from Gibco.

```

```

FEATURES             Location/Qualifiers
     source           1..289
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:238381"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Kid11"
                     /notes="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP Kid1;
                     prepared, and ss circles were made in vitro- Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Patima Bonaldo."

```

```

ORIGIN
Query Match      2.0%; Score 45; DB 1; Length 289;
Best Local Similarity 52.4%; Pred. No. 0.56;
Matches 99; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 85 GCCCGGAGCCCAAGCAACGACGAAAGCAGGACGACGGCGGGGTCTGTGCTTCCT 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 GACCCGACCGGGGAAAAGACTCAAAAGTAAACCTTTTCCATCCAGGAACCTGGA 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 145 GGCTACAAGTACCTCGGACCCCTTCAACGGACTCGACAAAGGGGAGCCCGCTCAACGGCG 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 TCTCTGGGTATTACCGGCGAGCAGCTACAGCCCCAGGGGTGGAGCCCGCTAGCCCTGTG 149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 205 GAGCGAGGGGCTCGAGCAGCAGCAAGGCTACGACCAAGCAGCTCAAAAGCGGTGACAAAT 264
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148 AGCTCACCAGTCTGACCACGACAAAGGGGCTCCCAAGCAGCTCGAAGAGCTCGACACAG 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 265 CCGTACCTG 273
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88 ACCCACCTG 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 13
AW206887/c
LOCUS      322 bp      mRNA      linear      EST 02-DEC-1999
DEFINITION UI-H-B11-afs-a-02-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
            IMAGE:272826 3', mRNA sequence.

```

```

ACCESSION AW206887
VERSION   AW206887.1 GI:6506383
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Homnidae; Homo.

```

```

REFERENCE 1 (bases 1 to 322)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Oligo-dT track not found. Not 1 site shown in beginning of sequence
           is likely internal to the message. cDNA library Preparation: M.B.
           Soares Lab Clone distribution: NCI-CGAP clone distribution
           information can be found through the I.M.A.G.E. Consortium/LLNL at:
           www-bio.llnl.gov/bbrp/image/image.html
           Seq primer: M13 Forward
           POLYA=No.

```

```

FEATURES             Location/Qualifiers
     source           1..322
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:272826"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Sub3"
                     /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site 1: Not 1; Site 2: Eco RI; The
                     NCI_CGAP Sub3 library is a subtracted library derived from
                     the NCI_CGAP Sub1 library, which is a subtracted library
                     derived from B1. B1 constitutes a mixture of 21
                     normalized or subtracted NCI_CGAP libraries:
                     NCI_CGAP Co4, NCI_CGAP Pr22, NCI_CGAP Pr28, NCI_CGAP_Co10,
                     NCI_CGAP Co16, NCI_CGAP Kid5, NCI_CGAP Kid12,
                     NCI_CGAP Kid3, NCI_CGAP Kid11, NCI_CGAP Lym2,
                     NCI_CGAP Br2, NCI_CGAP Co8, NCI_CGAP CLL1, NCI_CGAP_Le12,
                     NCI_CGAP_Brr23, NCI_CGAP_Lu5, NCI_CGAP_Lu24,
                     NCI_CGAP_Lu19, NCI_CGAP_Gc4, NCI_CGAP_Gc6,
                     NCI_CGAP_Brr25. These 21 libraries were pooled and a
                     single-stranded DNA preparation of the resulting mixture
                     was used as a tracer in a subtractive hybridization with
                     a driver whose composition is detailed below:
                     NCI_CGAP Kid3 pool 1 LLAM 3334-3337 3682-3683,
                     3798-3803 (IMAGE CloneIDs 1322376-1323911,
                     1456008-1456775, 1500552-1502855); NCI_CGAP Kid5 pool 1
                     LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs
                     1323912-1325831, 1471368-1472903, 1492104-1493255);
                     NCI_CGAP Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE
                     CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP GC4
                     pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE
                     CloneIDs 1257096-1258631, 1469064-1470983,

```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 19:15:02 ; Search time 314.186 Seconds  
(without alignments)  
12497.813 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_2223\_4431

Perfect score: 2209

Sequence: 1 atggctgcgagtggtatct.....ttacctaccgtccctgt 2209

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:\*

1: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*

2: /cgn2\_6/prodata/1/ina/5 COMB.seq.\*

3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*

4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*

5: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*

6: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*

7: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*

8: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*

9: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	2209	100.0	2211	3	US-09-807-802A-12
2	2209	100.0	4718	3	US-09-807-802A-1
3	2202.6	99.7	7447	3	US-10-216-870-11
4	2101.8	95.1	4693	3	US-09-807-802A-19
5	1798	81.4	1800	3	US-09-807-802A-14
6	1603	72.6	1605	3	US-09-807-802A-16
7	1458.6	66.0	4072	3	US-09-770-315-4
8	1458.6	66.0	4679	3	US-10-038-972A-12
9	1458.6	66.0	7557	3	US-09-770-315-3
10	1458.6	66.0	8698	3	US-09-770-315-2
11	1453.6	65.8	8179	3	US-09-438-268-5
12	1436.4	65.0	4680	2	US-08-254-358-1
13	1436.4	65.0	4680	2	US-08-475-391-1
14	1436.4	65.0	4680	2	US-08-709-609-1
15	1436.4	65.0	4680	6	PCT-US95-07178-1
16	1433.4	64.9	4681	3	US-09-807-802A-18
17	1408.2	63.7	4675	3	US-09-782-378A-1
18	1408.2	63.7	4675	3	US-09-782-378A-2
19	1241.8	56.2	8151	3	US-09-438-268-2
20	940	42.6	7744	3	US-10-216-870-14
21	933	42.2	7214	3	US-09-438-268-1
22	932	42.2	2208	3	US-09-532-594B-5
23	930.4	42.1	4767	3	US-09-532-594B-1
24	730.8	33.1	2307	3	US-09-533-427-7

25	730.8	33.1	4652	3	US-09-533-427-1	Sequence 1, Appli
26	727.8	32.9	2264	3	US-09-533-427-8	Sequence 8, Appli
27	727.8	32.9	2264	3	US-09-533-427-9	Sequence 9, Appli
28	627.2	28.4	1800	3	US-09-532-594B-17	Sequence 17, Appli
29	586.2	26.5	1617	3	US-09-532-594B-19	Sequence 19, Appli
30	441.6	20.0	2271	3	US-09-438-268-3	Sequence 3, Appli
31	59.2	2.7	5049	2	US-08-336-345-1	Sequence 1, Appli
32	59.2	2.7	5049	2	US-08-336-345-2	Sequence 2, Appli
33	59.2	2.7	5049	2	US-08-647-655-1	Sequence 1, Appli
34	59.2	2.7	5049	2	US-08-647-655-2	Sequence 2, Appli
35	50	2.3	969	3	US-09-807-802A-10	Sequence 10, Appli
36	50	2.3	1641	3	US-09-807-802A-6	Sequence 6, Appli
37	46	2.1	5511	3	US-08-928-361B-2	Sequence 2, Appli
38	46	2.1	5511	3	US-09-588-995A-2	Sequence 1, Appli
39	46	2.1	7334	3	US-08-928-361B-1	Sequence 1, Appli
40	46	2.1	7334	3	US-09-588-995A-1	Sequence 1, Appli
41	45	2.0	9053	3	US-09-976-594-306	Sequence 306, App
42	44.6	2.0	978	3	US-10-126-279-16	Sequence 16, Appli
43	44.6	2.0	978	3	US-10-286-606-16	Sequence 16, Appli
44	42.8	1.9	5163	3	US-08-700-651-1	Sequence 1, Appli
45	42.8	1.9	5163	3	US-08-928-361B-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-807-802A-12

; Sequence 12, Application US/09807802A

; Patent No. 6759237

; GENERAL INFORMATION:

; APPLICANT: Wilson, James M.

; APPLICANT: Xiao, Weidong

; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,

; TITLE OF INVENTION: Vectors and Host Cells Containing Same

; FILE REFERENCE: GNPVN.031USA

; CURRENT APPLICATION NUMBER: US/09/807,802A

; PRIOR FILING DATE: 2002-02-21

; PRIOR APPLICATION NUMBER: US 60/107,114

; PRIOR FILING DATE: 1998-11-05

; PRIOR APPLICATION NUMBER: PCT/US99/25694

; PRIOR FILING DATE: 1999-11-02

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 12

; LENGTH: 2211

; TYPE: DNA

; ORGANISM: AAV-1

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2208)

; OTHER INFORMATION:

US-09-807-802A-12

Query Match 100.0%; Score 2209; DB 3; Length 2211;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCTGCCGATGGTATCTTCAGATTGGCTCGAGCAACCTCTCTGAGGCGATTTCG	60
Db	1	ATGGCTGCCGATGGTATCTTCAGATTGGCTCGAGCAACCTCTCTGAGGCGATTTCG	60
QY	61	GAGTGTGGGACTTGAACCTTGAGCCCGAAGCCCAAGCCAAAGCCAGCAAGAGCAGGAC	120
Db	61	GAGTGTGGGACTTGAACCTTGAGCCCGAAGCCCAAGCCAAAGCCAGCAAGAGCAGGAC	120
QY	121	GACGGCGGGGCTCGTGTGCTTCTCGGTCTCAAGTACCTCGGACCCCTTCAACCGACTCGAC	180
Db	121	GACGGCGGGGCTCGTGTGCTTCTCGGTCTCAAGTACCTCGGACCCCTTCAACCGACTCGAC	180
QY	181	AAGGGGAGCCCGTCAACCGGGCGGACGAGCGGGCCCTCGAGCAGCAAGGCTTAGCAG	240
Db	181	AAGGGGAGCCCGTCAACCGGGCGGACGAGCGGGCCCTCGAGCAGCAAGGCTTAGCAG	240



QY 241 CAGAGCTCAAGCGGTTGACAAATCCGTACTGCGGTATAACCAAGCCGACGCGAGTTT 300  
DB 241 CAGAGCTCAAGCGGTTGACAAATCCGTACTGCGGTATAACCAAGCCGACGCGAGTTT 300  
QY 301 CAGGAGCGTCTCAAGAGAGATACGTTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360  
DB 301 CAGGAGCGTCTCAAGAGAGATACGTTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360  
QY 361 GCCAAGAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTTAAGACGGCTCT 420  
DB 361 GCCAAGAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTTAAGACGGCTCT 420  
QY 421 GGAAGAAACGTCGGTAGAGCAGTCGCCACAAGACAGACTCTCTCGGGGATCCGC 480  
DB 421 GGAAGAAACGTCGGTAGAGCAGTCGCCACAAGACAGACTCTCTCGGGGATCCGC 480  
QY 481 AAGACAGGCGCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAAGCTGGCAGCTCAGAG 540  
DB 481 AAGACAGGCGCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAAGCTGGCAGCTCAGAG 540  
QY 541 TCAGTCCCAGATCCAACTCTCGAGAACTTCAGAGAACTTCAGCAACCCCGCTCTGTGGACCT 600  
DB 541 TCAGTCCCAGATCCAACTCTCGAGAACTTCAGAGAACTTCAGCAACCCCGCTCTGTGGACCT 600  
QY 601 ACTACAATGGCTTCAGGCGGTGGCGACCAATGGCAGACAATAACGAAGCGCCGACGGA 660  
DB 601 ACTACAATGGCTTCAGGCGGTGGCGACCAATGGCAGACAATAACGAAGCGCCGACGGA 660  
QY 661 GTGGGTAAATGCTCAGGAAATTTGGCAATTCGCAATTCACATGGCTGGCGCAGAGTCAATC 720  
DB 661 GTGGGTAAATGCTCAGGAAATTTGGCAATTCGCAATTCACATGGCTGGCGCAGAGTCAATC 720  
QY 721 ACCACGACCGCGACCTTGGGCTTGGCCCACTCAATTAACCACTCTCAAGCAAAATC 780  
DB 721 ACCACGACCGCGACCTTGGGCTTGGCCCACTCAATTAACCACTCTCAAGCAAAATC 780  
QY 781 TCAGTGTCTCAAGCGGGGCGCAGCAACGACCACTACTTGGCTACAGCAACCCCGTGG 840  
DB 781 TCAGTGTCTCAAGCGGGGCGCAGCAACGACCACTACTTGGCTACAGCAACCCCGTGG 840  
QY 841 GGGTATTTTGAATTTCAACAGATTTCACTGCTCCACTTTTACACGACTGGAGCGACTC 900  
DB 841 GGGTATTTTGAATTTCAACAGATTTCACTGCTCCACTTTTACACGACTGGAGCGACTC 900  
QY 901 ATCAACAACAAATTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960  
DB 901 ATCAACAACAAATTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960  
QY 961 GTCNAGGAGGTACGACGAATGATGGCGTCACNACCATCGCTAAATTAACCTTACCAGCAG 1020  
DB 961 GTCNAGGAGGTACGACGAATGATGGCGTCACNACCATCGCTAAATTAACCTTACCAGCAG 1020  
QY 1021 GTTCAAGTCTTTCGAGCTCGAGTACAGCTTCCGTAAGTCTCGCTCGGCTCTGGCACCAG 1080  
DB 1021 GTTCAAGTCTTTCGAGCTCGAGTACAGCTTCCGTAAGTCTCGCTCGGCTCTGGCACCAG 1080  
QY 1081 GGTGCTCTCCCTCCGTTCCCGCGGACGTTTCATGATTCGCAATTCGCAATTCGCTACCTGACG 1140  
DB 1081 GGTGCTCTCCCTCCGTTCCCGCGGACGTTTCATGATTCGCAATTCGCAATTCGCTACCTGACG 1140  
QY 1141 CTCNAAATGGGAGCCAGCGGTGGAGCGTTGATCTCTTTTACTGCTGGAAATATTTCCCT 1200  
DB 1141 CTCNAAATGGGAGCCAGCGGTGGAGCGTTGATCTCTTTTACTGCTGGAAATATTTCCCT 1200  
QY 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTACCTTCAGCTTACACTTTGAGGAAGTGCCT 1260  
DB 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTACCTTCAGCTTACACTTTGAGGAAGTGCCT 1260  
QY 1261 TTCCACAGCAGCTACCGCAGCAGCCAGAGCCTGGACCGCTGTGATGTAATCTCTCATCGAC 1320  
DB 1261 TTCCACAGCAGCTACCGCAGCAGCCAGAGCCTGGACCGCTGTGATGTAATCTCTCATCGAC 1320

QY 1321 CAATACCTGTATTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAAGAC 1380  
DB 1321 CAATACCTGTATTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAAGAC 1380  
QY 1381 TTGCTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACCT 1440  
DB 1381 TTGCTGTTTAGCCGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAAACCTGGCTACCT 1440  
QY 1441 GGAACCTGTTATCCGACAGCGGCTTTCTAAAACAAAACAGACAAACAAACAGCAAT 1500  
DB 1441 GGAACCTGTTATCCGACAGCGGCTTTCTAAAACAAAACAGACAAACAAACAGCAAT 1500  
QY 1501 TTTTACTCGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCATATCAACCT 1560  
DB 1501 TTTTACTCGACTGGTCTTCAAAATATAACCTCAATGGCGGTGAATCATATCAACCT 1560  
QY 1561 GGCACCTGTCTGCGCTCACACAAAGACGACGAAAGTCTTTTCCCATGAGCGGTGTC 1620  
DB 1561 GGCACCTGTATGGCTCACACAAAGACGACGAAAGTCTTTTCCCATGAGCGGTGTC 1620  
QY 1621 ATGATTTTGGAAAGAGAGCGCGGAGCTTCAAAACACTGCAATTGGGCAATGTCTATGATT 1680  
DB 1621 ATGATTTTGGAAAGAGAGCGCGGAGCTTCAAAACACTGCAATTGGGCAATGTCTATGATT 1680  
QY 1681 ACAGACGAGAGGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740  
DB 1681 ACAGACGAGAGGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740  
QY 1741 GCAGTCAATTTCCAGAGCAGCAGCAACAGACCTGCGACCGGAGATGTGATGTGGGA 1800  
DB 1741 GCAGTCAATTTCCAGAGCAGCAGCAACAGACCTGCGACCGGAGATGTGATGTGGGA 1800  
QY 1801 GCATTAACCTGGCATGGTGTGCAAGATAGAGACGTGTACTGCAAGGTCCTATTTGGGCGC 1860  
DB 1801 GCATTAACCTGGCATGGTGTGCAAGATAGAGACGTGTACTGCAAGGTCCTATTTGGGCGC 1860  
QY 1861 AAAATTCCTCACACAGATGGACACTTTTCCCGCTCTCTTTATGGCGGCTTTGGACTC 1920  
DB 1861 AAAATTCCTCACACAGATGGACACTTTTCCCGCTCTCTTTATGGCGGCTTTGGACTC 1920  
QY 1921 AAGAACCCGCTCTCAGATCCTCATAAAAACAGCGCTTTTCTGCGAAATCCTCGGCG 1980  
DB 1921 AAGAACCCGCTCTCAGATCCTCATAAAAACAGCGCTTTTCTGCGAAATCCTCGGCG 1980  
QY 1981 GAGTTTTCAGCTAAGAGTTTGTCTTCAATTCATCACCCTCAATCTCAGAGCAAGTGAAT 2040  
DB 1981 GAGTTTTCAGCTAAGAGTTTGTCTTCAATTCATCACCCTCAATCTCAGAGCAAGTGAAT 2040  
QY 2041 GTGGAATTTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCGAAGTGCAG 2100  
DB 2041 GTGGAATTTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCGAAGTGCAG 2100  
QY 2101 TACACATCCAAATTTGCAAAATCTGCCAAGCTTTGATTTTACTGTGGACAAATGGACTT 2160  
DB 2101 TACACATCCAAATTTGCAAAATCTGCCAAGCTTTGATTTTACTGTGGACAAATGGACTT 2160  
QY 2161 TATATGAGCCTCGGCCCATTTGGACCCGCTTACCTTACCTGTCCTCCCTGT 2209  
DB 2161 TATATGAGCCTCGGCCCATTTGGACCCGCTTACCTTACCTGTCCTCCCTGT 2209

## RESULT 2

US-09-807-802A-1  
; Sequence 1, Application US/09807802A  
; Patent No. 6759237  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,  
; TITLE OF INVENTION: Vectors and Host Cells Containing Same  
; FILE REFERENCE: GNPVN.031USA  
; CURRENT APPLICATION NUMBER: US/09/807,802A  
; CURRENT FILING DATE: 2002-02-21

```

; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4718
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (335)..(2206)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2223)..(4430)
; OTHER INFORMATION:
; US-09-807-802A-1

```

	Query Match	100.0%	Score 2209;	DB 3;	Length 4718;
	Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Gaps 0;
	Matches 2209;	Conservative 0;			
Qy	1	ATGCTCCGATGGTTATCTTCAGATTGGCTCGAGACAACTCTCTGAGGGCATTCGC	60		
Db	2223	ATGCTCCGATGGTTATCTTCAGATTGGCTCGAGACAACTCTCTGAGGGCATTCGC	2282		
Qy	61	GAGTGTGGGACTTGAAACCCTGGAGCCCCGAAGCCCACAAGCCAACCAAGCAAGCAGAC	120		
Db	2283	GAGTGTGGGACTTGAAACCCTGGAGCCCCGAAGCCCACAAGCCAACCAAGCAAGCAGAC	2342		
Qy	121	GACGGCCGGGCTCTGGTGCTTCCTGGGTACAAGTAACCTCGGACCCCTTCAACCGACTCGAC	180		
Db	2343	GACGGCCGGGCTCTGGTGCTTCCTGGGTACAAGTAACCTCGGACCCCTTCAACCGACTCGAC	2402		
Qy	181	AAGGGGAGCCCGTCAACGCGCGGAGCGCACGCCCTTCGAGCAACCAAGCCCTACGAC	240		
Db	2403	AAGGGGAGCCCGTCAACGCGCGGAGCGCACGCCCTTCGAGCAACCAAGCCCTACGAC	2462		
Qy	241	CAGCAGCTCAAAGCGGGTGACATCTCGTACTTGGCGGTATAACAGCGCGAAGCCGAGTTTT	300		
Db	2463	CAGCAGCTCAAAGCGGGTGACATCTCGTACTTGGCGGTATAACAGCGCGAAGCCGAGTTTT	2522		
Qy	301	CAGGAGCGTCTGCAAGAAGATACCTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCAG	360		
Db	2523	CAGGAGCGTCTGCAAGAAGATACCTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCAG	2582		
Qy	361	GCCAAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCCT	420		
Db	2583	GCCAAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCCT	2642		
Qy	421	GGAAGAAAACCTCCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCTCGGGCATCGGC	480		
Db	2643	GGAAGAAAACCTCCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCTCGGGCATCGGC	2702		
Qy	481	AAGACAGGCCAGACCCCGCTAAAAAGAGACTCAATTTTGTCTAGA CTGGCGACTCAGAG	540		
Db	2703	AAGACAGGCCAGACCCCGCTAAAAAGAGACTCAATTTTGTCTAGA CTGGCGACTCAGAG	2762		
Qy	541	TCAGTCCCGATGCACAACTCTCGAGAAACCTTCAGCAACCCCGCGTCTGTGGGACCT	600		
Db	2763	TCAGTCCCGATGCACAACTCTCGAGAAACCTTCAGCAACCCCGCGTCTGTGGGACCT	2822		
Qy	601	ACTACATGGCTTCAGGCGGTGGCGACCAATGGCAGACAATNACGAAGGGCCGACCGGA	660		
Db	2823	ACTACATGGCTTCAGGCGGTGGCGACCAATGGCAGACAATNACGAAGGGCCGACCGGA	2882		
Qy	661	GTGGTAAATGCTCAGGAAAATTGGCATTCGATTCCACATGGCTGGGCGACAGAGTCAATC	720		
Db	2883	GTGGTAAATGCTCAGGAAAATTGGCATTCGATTCCACATGGCTGGGCGACAGAGTCAATC	2942		
Qy	721	ACCACCGACACCCGACCTGGGCGCTTGCCCACTACAAATAACCACTCTACAAAGCAAAATC	780		

Db	2943	ACCAACAGACCCGACCTGGGCTTTGCCACCTA	CAATAAACCACTCTACAGCAATC	3002
Qy	781	TCAGTGTCTTAAACGGGGGACGAAACGAAACCA	CTACTTGGGCTTACAGCACCCCTGG	840
Db	3003	TCAGTGTCTTAAACGGGGGACGAAACGAAACCA	CTACTTGGGCTTACAGCACCCCTGG	3062
Qy	841	GGGTATTTTGAATTTCAACAGATTTCCATGCGCA	CTTTTCAACAGTGTGGCAGCGACTC	900
Db	3063	GGGTATTTTGAATTTCAACAGATTTCCATGCGCA	CTTTTCAACAGTGTGGCAGCGACTC	3122
Qy	901	ATCAACAAATTTGGGGATTTCCGGCCCAAGAGACT	CAAACTTTCAAACCTTTCAAACATCCAA	960
Db	3123	ATCAACAAATTTGGGGATTTCCGGCCCAAGAGACT	CAAACTTTCAAACCTTTCAAACATCCAA	3182
Qy	961	GTCAAGAGAGTCAACGAAATGATGGGCTCACAA	CCATCGCTAATAAACCTTACAGCAGC	1020
Db	3183	GTCAAGAGAGTCAACGAAATGATGGGCTCACAA	CCATCGCTAATAAACCTTACAGCAGC	3242
Qy	1021	GTTCNAAGTCTTTCGGACTCGGAGTACCAAGCTT	CCGTACGTCTCGGCTCTGGCGACCA	1080
Db	3243	GTTCNAAGTCTTTCGGACTCGGAGTACCAAGCTT	CCGTACGTCTCGGCTCTGGCGACCA	3302
Qy	1081	GGCTGCCCTCCCTCCGTTTCCGGCGGACGTGTT	CATGATTCGCAATACGGCTACCTGACG	1140
Db	3303	GGCTGCCCTCCCTCCGTTTCCGGCGGACGTGTT	CATGATTCGCAATACGGCTACCTGACG	3362
Qy	1141	CTCAACAAATGGCAGCCAGCGTGGGACGTTTCAT	TCCTTTTACTGCCTTGGGAATATTCCCT	1200
Db	3363	CTCAACAAATGGCAGCCAGCGTGGGACGTTTCAT	TCCTTTTACTGCCTTGGGAATATTCCCT	3422
Qy	1201	TCTCAGATGCTGAGAACGGGCAACAACTTTTAC	CTTTCAGCTTACGCTTTGAGGAAGTCCT	1260
Db	3423	TCTCAGATGCTGAGAACGGGCAACAACTTTTAC	CTTTCAGCTTACGCTTTGAGGAAGTCCT	3482
Qy	1261	TTCCACAGCAGCTACGCGACACGACGAGCGCT	CGGCTGATGTAATCCTCTCATCGAC	1320
Db	3483	TTCCACAGCAGCTACGCGACACGACGAGCGCT	CGGCTGATGTAATCCTCTCATCGAC	3542
Qy	1321	CAATACCTGTATTAACCTGAAACAGAACTCAAA	ATCAGTCCGGAAGTCCCAAAACAAGGAC	1380
Db	3543	CAATACCTGTATTAACCTGAAACAGAACTCAAA	ATCAGTCCGGAAGTCCCAAAACAAGGAC	3602
Qy	1381	TTGCTGTTTATGCGTGGGTCTCAGCTGGGATGT	CTGTTTTCAGCCCAAAACTGGCTACCT	1440
Db	3603	TTGCTGTTTATGCGTGGGTCTCAGCTGGGATGT	CTGTTTTCAGCCCAAAACTGGCTACCT	3662
Qy	1441	GGACCTGTATTCGGCAGCAGCGGTTCTTAAAA	CAAAAAACAGACAAACAACAGCAAT	1500
Db	3663	GGACCTGTATTCGGCAGCAGCGGTTCTTAAAA	CAAAAAACAGACAAACAACAGCAAT	3722
Qy	1501	TTTACCTGGACTGGTGTCTTCAAAATATAACCT	CTCAATGGGCGTGAATCCATCATCAACCT	1560
Db	3723	TTTACCTGGACTGGTGTCTTCAAAATATAACCT	CTCAATGGGCGTGAATCCATCATCAACCT	3782
Qy	1561	GGCACTGCTATGGCCTCACAAAGACGACGAAGA	CAAGTGTCTTTTCCCATGAGCGGTGTC	1620
Db	3783	GGCACTGCTATGGCCTCACAAAGACGACGAAGA	CAAGTGTCTTTTCCCATGAGCGGTGTC	3842
Qy	1621	ATGATTTTTTGGAAAAGAGCGCGGAGCTTCAAA	CACTCANTTGGACAATGTATCATATT	1680
Db	3843	ATGATTTTTTGGAAAAGAGCGCGGAGCTTCAAA	CACTCANTTGGACAATGTATCATATT	3902
Qy	1681	ACAGCAAGAGGAAATTTAAAGCCACCTTAAAC	CTGTGGCCACCGAAAAGATTTGGACCGGTG	1740
Db	3903	ACAGCAAGAGGAAATTTAAAGCCACCTTAAAC	CTGTGTGGCCACCGAAAAGATTTGGACCGGTG	3962
Qy	1741	GCAGTCAATTTCCAGACGACGACACAGACCTTC	GCGACCGGAGATGTGCATGTATGGGA	1800
Db	3963	GCAGTCAATTTCCAGACGACGACACAGACCTTC	GCGACCGGAGATGTGCATGTATGGGA	4022
Qy	1801	GCATTAACCTGGCATGTGTGGCAAGATAGAGAC	GTGTACCTGCGAGGTCCCATTTGGGCC	1860

Db 4023 GCATTACCTGGCATGCTGGCAAGATAGAGAGCTGTACCTGCAGGGTCCCATTTTGGGCG 4082  
QY 1861 AAAATTCTCACAAGATGGACATTTTCAACCGTCTCTCTTATATGGGCGGCTTTGGACTC 1920  
Db 4083 AAAATTCTCACAAGATGGACATTTTCAACCGTCTCTCTTATATGGGCGGCTTTGGACTC 4142  
QY 1921 AAGAACCCGCTTCTCAGATCTCATCAAAACACGCGTTCCTTGGGAATCCTCCGGG 1980  
Db 4143 AAGAACCCGCTTCTCAGATCTCATCAAAACACGCGTTCCTTGGGAATCCTCCGGG 4202  
QY 1981 GAGTTTTAGCTACAAAGTTTGTCTTATTCATCACCCTTACTCCACAGGACAAAGTGA 2040  
Db 4203 GAGTTTTAGCTACAAAGTTTGTCTTATTCATCACCCTTACTCCACAGGACAAAGTGA 4262  
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTG 2100  
Db 4263 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTG 4322  
QY 2101 TACACATCCAATATGCAAAATCTGCAACGTTGATTTTACTGTGGACAAACATGGACTT 2160  
Db 4323 TACACATCCAATATGCAAAATCTGCAACGTTGATTTTACTGTGGACAAACATGGACTT 4382  
QY 2161 TATACCTGAGCTCGCCCCATTTGGCACCCGTTACCTTACCCGTCCTCGT 2209  
Db 4383 TATACCTGAGCTCGCCCCATTTGGCACCCGTTACCTTACCCGTCCTCGT 4431

## RESULT 3

US-10-216-870-11  
; Sequence 11, Application US/10216870  
; Patent No.: 6723551  
; GENERAL INFORMATION:  
; APPLICANT: KOTIN, ROBERT M  
; APPLICANT: URABE, MASASHI  
; APPLICANT: DING, CHUAN-TIAN  
; TITLE OF INVENTION: PRODUCTION OF ADENO-ASSOCIATED VIRUS IN INSECT CELLS  
; FILE REFERENCE: 402133  
; CURRENT APPLICATION NUMBER: US/10/216,870  
; CURRENT FILING DATE: 2002-08-13  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 11  
; LENGTH: 7447  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-216-870-11

Query Match 99.7%; Score 2202.6; DB 3; Length 7447;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2205; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 ATGGCTGCCGATGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGSCATTGCG 60  
Db 4621 ACGGCTGCCGACGGTTATCTACCCGATTGGCTCGAGGACAACTCTCTGAGGSCATTGCG 4680  
QY 61 GAGTGGTGGGACTTGAACCTTGAGCGCCGAGAGCCCAAGCCAAAGCCAAAGCAGGAG 120  
Db 4681 GAGTGGTGGGACTTGAACCTTGAGCGCCGAGAGCCCAAGCCAAAGCAGGAG 4740  
QY 121 GACGGCGGGGTCTGGTCTCTTCTGCTACAACTGTAAGTACTCGGACCTTCAACGACTCG 180  
Db 4741 GACGGCGGGGTCTGGTCTCTTCTGCTACAACTGTAAGTACTCGGACCTTCAACGACTCG 4800  
QY 181 AAGGGGAGCCCTCAACGCGGCGGAGCGGCGCTCGAGCAACGAGGCTACGAC 240  
Db 4801 AAGGGGAGCCCTCAACGCGGCGGAGCGGCGCTCGAGCAACGAGGCTACGAC 4860  
QY 241 CAGCAGCTCAAGCGGGTGAACAATCCGTACTCTGCGGTATAACCAACGCGGCGGAGTTT 300  
Db 4861 CAGCAGCTCAAGCGGGTGAACAATCCGTACTCTGCGGTATAACCAACGCGGCGGAGTTT 4920

QY 301 CAGGAGCGTCTGCAAGAGATACGTTCTTTGGGGCAACCTCGGCGGAGCAGTCTTCCAG 360  
Db 4921 CAGGAGCGTCTGCAAGAGATACGTTCTTTGGGGCAACCTCGGCGGAGCAGTCTTCCAG 4980  
QY 361 GCCAAGAACGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTTAAGACGGCTCT 420  
Db 4981 GCCAAGAACGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTTAAGACGGCTCT 5040  
QY 421 GGAAGAAACGTCGCTGAGAGAGTGCACCAAGAGCCAGACTCTCTCTCGGCGATCGGC 480  
Db 5041 GGAAGAAACGTCGCTGAGAGAGTGCACCAAGAGCCAGACTCTCTCTCGGCGATCGGC 5100  
QY 481 AAGCAGCGCAGCAGCGCTTAAAGAGAGCTCAATTTTGGTTCAGACTGGGAGCTCAGAG 540  
Db 5101 AAGCAGCGCAGCAGCGCTTAAAGAGAGCTCAATTTTGGTTCAGACTGGGAGCTCAGAG 5160  
QY 541 TCAGTCCCGGATCCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGGACCT 600  
Db 5161 TCAGTCCCGGATCCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGGACCT 5220  
QY 601 ACTACAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATTAAGAAAGCGCGCAGCA 660  
Db 5221 ACTACAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATTAAGAAAGCGCGCAGCA 5280  
QY 661 GTGGTAAATGCTTCAGGAAATTTGGCAATTCGATTTCCACATGGCTGGGCGAGAGTCA 720  
Db 5281 GTGGTAAATGCTTCAGGAAATTTGGCAATTTCCATGTCATGGCTGGGCGAGAGTCA 5340  
QY 721 ACCACGACGCCGACCTGGGCGCTTGGCCACCTTACAAATAACCACTCTTACAAGCAAA 780  
Db 5341 ACCACGACGCCGACCTGGGCGCTTGGCCACCTTACAAATAACCACTCTTACAAGCAAA 5400  
QY 781 TCCAGTGTTCACCGGGGGCGAGCAACCAACCACTACTTCCGCTACAGCACCCCTGG 840  
Db 5401 TCCAGTGTTCACCGGGGGCGAGCAACCAACCACTACTTCCGCTACAGCACCCCTGG 5460  
QY 841 GGGTAAATTTGATTTCAACAGATTTCACCTGCGACCTTTTCCACGCTGACTGGCAGGACTC 900  
Db 5461 GGGTAAATTTGATTTCAACAGATTTCACCTGCGACCTTTTCCACGCTGACTGGCAGGACTC 5520  
QY 901 ATCAACAAACAAATTTGGGGATTCCGGGCCAAGAGACTCAACTTTCAAACTCTTCAACATCAA 960  
Db 5521 ATCAACAAACAAATTTGGGGATTCCGGGCCAAGAGACTCAACTTTCAAACTCTTCAACATCAA 5580  
QY 961 GTCAAGAGGTCACGACGAATGATGGCTGACCAACCACTGCTGCTTACAGCAGCG 1020  
Db 5581 GTCAAGAGGTCACGACGAATGATGGCTGACCAACCACTGCTGCTTACAGCAGCG 5640  
QY 1021 GTTCAAGTCTTCTCGGACTCCGAGTACAGCTTCCGTTAGTCTCGGCTCTGCGCACCGAG 1080  
Db 5641 GTTCAAGTCTTCTCGGACTCCGAGTACAGCTTCCGTTAGTCTCGGCTCTGCGCACCGAG 5700  
QY 1081 GGCTGCTCTCTCCGTTCCGGCGGACGTTGTTTATGATTCCGCAATACGGCTACCTGAGC 1140  
Db 5701 GGCTGCTCTCTCCGTTCCGGCGGACGTTGTTTATGATTCCGCAATACGGCTACCTGAGC 5760  
QY 1141 CTCAA CAATGGCAGCAAGCGCTGGAAGCTTCACTTTTACTGCTGGAATATTTCCCT 1200  
Db 5761 CTCAA CAATGGCAGCAAGCGCTGGAAGCTTCACTTTTACTGCTGGAATATTTCCCT 5820  
QY 1201 TCTCAGATGCTGAGAACGGGCAACCTTTACCTTACGCTACACCTTTGAGGAGTGCCT 1260  
Db 5821 TCTCAGATGCTGAGAACGGGCAACCTTTACCTTACGCTACACCTTTGAGGAGTGCCT 5880  
QY 1261 TTTCCACAGCAGCTACGCGCACAGCAGCCTGGAACCGGCTGATGAATCTCTCTCATTCGAC 1320  
Db 5881 TTTCCACAGCAGCTACGCGCACAGCAGCCTGGAACCGGCTGATGAATCTCTCTCATTCGAC 5940  
QY 1321 CAATACCTGTATTACTGAAAGAACTCAAAATTCAGTCCGGAAGTGCCTCAAAACAGAGAC 1380  
Db 5941 CAATACCTGTATTACTGAAAGAACTCAAAATTCAGTCCGGAAGTGCCTCAAAACAGAGAC 6000  
QY 1381 TTGCTGTTTAGCGGTGCTCCAGCTGGCAATGCTGTTTACGCCCAAAACTGGGTACCT 1440

```

Db      6001  |||||TTTGCTGTTTATGCGTGGGCTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAAACTGGCTACCT 6060
Qy      1441  GGACCCCTGTTATCGGCAGCGCGTCTTCTAAACCAAAACAGACAAACAGCAAT 1500
Db      6061  GGACCCCTGTTATCGGCAGCGCGTCTTCTAAACCAAAACAGACAAACAGCAAT 6120
Qy      1501  TTTTACCTGGACTGGTCTTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCT 1560
Db      6121  TTTTACCTGGACTGGTCTTTCAAAATATAACCTCAATGGCGGTGAATCCATCATCAACCT 6180
Qy      1561  GGCACTGCTATGCGCTCAACAAGAAGAGCAAGCAAGTCTTTCCCATGAGCGGTGTC 1620
Db      6181  GGCACTGCTATGCGCTCAACAAGAAGAGCAAGCAAGTCTTTCCCATGAGCGGTGTC 6240
Qy      1621  ATGATTTTTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATGCAACAATGTCATGATT 1680
Db      6241  ATGATTTTTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCAATGCAACAATGTCATGATT 6300
Qy      1681  ACAGACGAAGAGAAAATTAAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740
Db      6301  ACAGACGAAGAGAAAATTAAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 6360
Qy      1741  GCAGTCAATTTCCAGAGCAGCAGCAAGACCCCTGCGACCGGAGATGTGATGCTATGGGA 1800
Db      6361  GCAGTCAATTTCCAGAGCAGCAGCAAGACCCCTGCGACCGGAGATGTGATGCTATGGGA 6420
Qy      1801  GCATTAACCTGGCATGTGTGGCAAGATAGAGAGCTGCTACTGACAGGTCCTCAATTTGGGCC 1860
Db      6421  GCATTAACCTGGCATGTGTGGCAAGATAGAGAGCTGCTACTGACAGGTCCTCAATTTGGGCC 6480
Qy      1861  AAAATTCCTCACACAGATGGACACTTTTCAACCCCTCTCTCTTATGGCGGCTTTGGACTC 1920
Db      6481  AAAATTCCTCACACAGATGGACACTTTTCAACCCCTCTCTCTTATGGCGGCTTTGGACTC 6540
Qy      1921  AGAAACCGGCTCTCAGATCTCATCAAAACACACGCTGTTCTCTGGGAATCTCTCGGG 1980
Db      6541  AGAAACCGGCTCTCAGATCTCATCAAAACACACGCTGTTCTCTGGGAATCTCTCGGG 6600
Qy      1981  GAGTTTTCAGCTACAAGTTTGCTTCATTCATCAACCCATACCTCCACAGACAAGTGAGT 2040
Db      6601  GAGTTTTCAGCTACAAGTTTGCTTCATTCATCAACCCATACCTCCACAGACAAGTGAGT 6660
Qy      2041  GTGGAATTAATGAGAGCTGCAAGAAAGAAACAGACGCTGGAATCCGGAATGTCAG 2100
Db      6661  GTGGAATTAATGAGAGCTGCAAGAAAGAAACAGACGCTGGAATCCGGAATGTCAG 6720
Qy      2101  TACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGACAAACAATGACTT 2160
Db      6721  TACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGACAAACAATGACTT 6780
Qy      2161  TATACGTAGCTCGCCCATTTGGCACCGCTTACCTTACCGTCCCTGT 2209
Db      6781  TATACGTAGCTCGCCCATTTGGCACCGCTTACCTTACCGTCCCTGT 6829
```

```

RESULT 4
US-09-807-802A-19
; Sequence 19, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; FILE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 4683
; TYPE: DNA
; ORGANISM: RAV-6
US-09-807-802A-19

Query Match      95.1%; Score 2101.8; DB 3; Length 4683;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2142; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

Qy      1  ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGCAAACTCTCTGAGGCGATTCCG 60
Db      2208  ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGCAAACTCTCTGAGGCGATTCCG 2267
Qy      61  GAGTGTGGGACTTCAAACTCGAGCCCGAAAGCCCAAGCCAAACAGCAAAAGCAGGAC 120
Db      2268  GAGTGTGGGACTTCAAACTCGAGCCCGAAAGCCCAAGCCAAACAGCAAAAGCAGGAC 2327
Qy      121  GACGCGCGGGTCTGGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC 180
Db      2328  GACGCGCGGGTCTGGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC 2387
Qy      181  AAGGGGGAGCCCGTCAACGCGGGGACGCGAGCGGCTCGAGCAGCAAGGCTTACGAC 240
Db      2388  AAGGGGGAGCCCGTCAACGCGGGGACGCGAGCGGCTCGAGCAGCAAGGCTTACGAC 2447
Qy      241  CAGCAGCTCAAAAGCGGGTGACAATCCGTACTCGGTATAACCAACCGCAGCGAGTTT 300
Db      2448  CAGCAGCTCAAAAGCGGGTGACAATCCGTACTCGGTATAACCAACCGCAGCGAGTTT 2507
Qy      301  CAGGAGCGTCTCAAGAAAGATACGCTTTTGGGGGCAACTCGGGGAGAGTCTTCCAG 360
Db      2508  CAGGAGCGTCTCAAGAAAGATACGCTTTTGGGGGCAACTCGGGGAGAGTCTTCCAG 2567
Qy      361  GCCAAGAACGGGTTCTCGAACTCTCGCTCTGGTTGAGGAGGCGCTTAAGCGGCTCCT 420
Db      2568  GCCAAGAACGGGTTCTCGAACTTTTGGTCTGGTTGAGGAGGTTGCTAAGCGGCTCCT 2627
Qy      421  GGAAAGAAACGCTCCGTTAGAGCAGTCCCAAGAGCCAGACTCTCTCTCGGCGATCGGC 480
Db      2628  GGAAAGAAACGCTCCGTTAGAGCAGTCCCAAGAGCCAGACTCTCTCTCGGCGATCGGC 2687
Qy      481  AAGACAGGCGCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTCAAGTGGCGACTCAGAG 540
Db      2688  AAGACAGGCGCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTCAAGTGGCGACTCAGAG 2747
Qy      541  TCAGTCCCGCATTCACAACCTCTCGAGAACTTCCAGAGAACCCCGCTGCTGTGGACCT 600
Db      2748  TCAGTCCCGCATTCACAACCTCTCGAGAACTTCCAGAGAACCCCGCTGCTGTGGACCT 2807
Qy      601  ACTTACAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATAAAGAAAGCGCGCAGGA 660
Db      2808  ACTTACAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATAAAGAAAGCGCGCAGGA 2867
Qy      661  GTGGGTAATGCTCTCAGGAAATTTGGCATTTCCATGTTGGTGGCGCAGAGTCAATC 720
Db      2868  GTGGGTAATGCTCTCAGGAAATTTGGCATTTCCATGTTGGTGGCGCAGAGTCAATC 2927
Qy      721  ACCACAGACCCCGCACTTGGGCTTGGCCACCTACAATACCACTCTTACAGCAATC 780
Db      2928  ACCACAGACCCCGCACTTGGGCTTGGCCACCTACAATACCACTCTTACAGCAATC 2987
Qy      781  TCCAGTGTCTTCAACGGGGCGCAGCAACCACTACTTCCGCTTACAGCAACCCCTCG 840
Db      2988  TCCAGTGTCTTCAACGGGGCGCAGCAACCACTACTTCCGCTTACAGCAACCCCTCG 3047
Qy      841  GGGTATTTTGAATTTCAACAGATTTCCACTGCGACTTTTTCACCAAGTGTGGCAGCAATC 900
Db      3048  GGGTATTTTGAATTTCAACAGATTTCCACTGCGACTTTTTCACCAAGTGTGGCAGCAATC 3107
Qy      901  ATCAACAACAATTTGGGATTTCCGCGCCCAAGAGACTCAACTTCAACTTCTTCAACATCAA 960
Db      |||||
```

Db 3108 ATCAACAAATTTGGGATTTCGGCCCAAGAGACTCAACTTCAAGCTCTTCAACATCOAA 3167  
QY 961 GTCAAGGAGGTACAGACGAATGATGGCTGTCACAAACATCGCTAAATAACCTTACGACGAG 1020  
Db 3168 GTCAAGGAGGTACAGACGAATGATGGCTGTCACAGCAATCGCTAAATAACCTTACGACGAG 3227  
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTGCTGCTCGGCTCTCGGACACAG 1080  
Db 3228 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTGCTGCTCGGCTCTCGGACACAG 3287  
QY 1081 GGCTGCTCCCTCCGTTCCGGCGGACAGCTGTTTCAATGATTCGCAATACGCTACCTGAGC 1140  
Db 3288 GGCTGCTCCCTCCGTTCCGGCGGACAGCTGTTTCAATGATTCGCAATACGCTACCTGAGC 3347  
QY 1141 CTCACAAATGGGAGCCCAAGCCGTGGGACGTTTCACTTTTACTGCTTGAATATTTTCCCT 1200  
Db 3348 CTCACAAATGGGAGCCCAAGCCGTGGGACGCTCATCTTTTACTGCTTGAATATTTTCCCA 3407  
QY 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTTGAGGAAGTGCT 1260  
Db 3408 TGCAGATGCTGAGAACGGGCAATTAATTTTACTTCAGCTACACCTTCGAGGACGTGCT 3467  
QY 1261 TTCCACAGAGCTACGCGACAGCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGAC 1320  
Db 3468 TTCCACAGAGCTACGCGACAGCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGAC 3527  
QY 1321 CAATACCTGTATTTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAGGAC 1380  
Db 3528 CAGTACCTGTATTTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAGGAC 3587  
QY 1381 TTGCTGTTTGTAGCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACCTGGCTACCT 1440  
Db 3588 TTGCTGTTTGTAGCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAACCTGGCTACCT 3647  
QY 1441 GGACCTGTATTCGGCAGCAGCGCGTTTCTAAACAAACAGACAAACAAACAGCAAT 1500  
Db 3648 GGACCTGTATTCGGCAGCAGCGCGTTTCTAAACAAACAGACAAACAAACAGCAAT 3707  
QY 1501 TTTTACCTGAGCTGGTCTTCAAAATATAACCTCAATGCGGCTGGAATTCATCATCAACCT 1560  
Db 3708 TTTTACCTGAGCTGGTCTTCAAAATATAACCTTAATGGCGGTGAATCTATAATCAACCT 3767  
QY 1561 GGCACCTGTATGGCTTCACAAAGACGAGCAAGCAAGTCTTTTCCATGAGCGGTGTC 1620  
Db 3768 GGCACCTGTATGGCTTCACAAAGACGAGCAAGCAAGTCTTTTCCATGAGCGGTGTC 3827  
QY 1621 ATGATTTTTGGAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGACAATGTCAATGAT 1680  
Db 3828 ATGATTTTTGGAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGACAATGTCAATGATC 3887  
QY 1681 ACAGACGAAGAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740  
Db 3888 ACAGACGAAGAGAGAAATCAAGGCCACTAAACCCGTGGCCACCGAAAGATTTGGGACGTG 3947  
QY 1741 CGAGTCAATTTTCAGAGCAGCAGACAGACCTTCGACCGGAGATGTGATGCTATGGGA 1800  
Db 3948 CGAGTCAATTTTCAGAGCAGCAGACAGACCTTCGACCGGAGATGTGATGCTATGGGA 4007  
QY 1801 GCATTAACCTGGCATGTGTGGCAAGATAGAGACGTGTACCTGACGGGTGCCAATTTGGGCG 1860  
Db 4008 GCCTTAACCTGGCATGTGTGGCAAGATAGAGACGTGTACCTGACGGGTGCCAATTTGGGCG 4067  
QY 1861 AAAATTCCTCAGACAGATGGACACTTTTCAACCGTCTCTCTTATGGGGCGCTTTGGACTC 1920  
Db 4068 AAAATTCCTCAGACAGATGGACACTTTTCAACCGTCTCTCTCTATGGGGCGCTTTGGACT 4127  
QY 1921 AGAAGCCGCTCCTCAGATCCTCATCAAAACACAGCCCTGTCTTCGCAATCCTCCGGCG 1980  
Db 4128 AGAAGCCGCTCCTCAGATCCTCATCAAAACACAGCCCTGTCTTCGCAATCCTCCGGCG 4187  
QY 1981 GAGTTTTTCAGCTACAAAGTTTGTCTTCAATTCATCACCCTTCTCCACAGGACAAGTGAGT 2040  
Db 4188 GAGTTTTTCAGCTACAAAGTTTGTCTTCAATTCATCACCCTTCTCCACAGGACAAGTGAGC 4247

QY 2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGACGAGCGCTGGAATCCCGAAGTGCAG 2100  
Db 4248 GTGGAGATTTGAATGGGAGCTGCAGAAAGAAACAGACGAGCGCTGGAATCCCGAAGTGCAG 4307  
QY 2101 TACACATCAATTTATGAAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAAATGGACTT 2160  
Db 4308 TATACATCTAATTTATGAAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAAATGGACTT 4367  
QY 2161 TATACTAGCCTCGCCCCCATTTGGCACCCGTTTACCTTACCCCTGCCCTGT 2209  
Db 4368 TATACTAGCCTCGCCCCCATTTGGCACCCGTTTACCTTACCCCTGCCCTGT 4416

## RESULT 5

US-09-807-802A-14  
; Sequence 14, Application US/09807802A  
; Patent No. 6759237  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: Adeno-Associated Virus Serotype 1 Nucleic Acid Sequences,  
; FILE OF INVENTION: Vectors and Host Cells Containing Same  
; FILE REFERENCE: GNPVN.03105A  
; CURRENT APPLICATION NUMBER: US/09/807,802A  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/107,114  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: PCT/US99/25694  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 1800  
; TYPE: DNA  
; ORGANISM: AAV-1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1797)  
; OTHER INFORMATION:  
US-09-807-802A-14

Query Match 81.4%; Score 1798; DB 3; Length 1800;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1798; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 412 ACGGCTCTCTGGAAGAAACGTCGCTAGAGCAGTCGCCACAGAGCCAGACTCTCTCTCG 471  
Db 1 ACGGCTCTCTGGAAGAAACGTCGCTAGAGCAGTCGCCACAGAGCCAGACTCTCTCTCG 60  
QY 472 GGCATCGGCAAGACAGAGCCGAGCCGCTAAAAAGAGACTCAATTTTGGTCAGACTGGC 531  
Db 61 GGCATCGGCAAGACAGAGCCGAGCCGCTAAAAAGAGACTCAATTTTGGTCAGACTGGC 120  
QY 532 GACTCAGAGTCACTCCCGATCCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTGT 591  
Db 121 GACTCAGAGTCACTCCCGATCCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTGT 180  
QY 592 GTGGACCTTACTTCAATGGCTTCAGCGGTGGCGCAACCAATGGCAGACAAATACGAAGC 651  
Db 181 GTGGACCTTACTTCAATGGCTTCAGCGGTGGCGCAACCAATGGCAGACAAATACGAAGC 240  
QY 652 GCCAGGAGTGGGTAATGCTCAGGAAATTTGGGATTTGGGATTCACATGGCTGGCGAC 711  
Db 241 GCCAGGAGTGGGTAATGCTCAGGAAATTTGGGATTTGGGATTCACATGGCTGGCGAC 300  
QY 712 AGAGTCATCACCACAGCAGCCGACCTGGGGCTTGGCCACCTACAAATACCACTCTAC 771  
Db 301 AGAGTCATCACCACAGCAGCCGACCTGGGGCTTGGCCACCTACAAATACCACTCTAC 360  
QY 772 AAGCAAAATCTCAGTGTCTTCAACGGGGGCGCAGCAACCACTACTTTCGGCTACAGC 831  
Db 361 AAGCAAAATCTCAGTGTCTTCAACGGGGGCGCAGCAACCACTACTTTCGGCTACAGC 420



QY 832 ACCCCCTGGGGTATTTTGAATTTCAACAGATTCACCTGCGCACTTTTCAACAGTGACTGG 891  
DB 421 ACCCCCTGGGGTATTTTGAATTTCAACAGATTCACCTGCGCACTTTTCAACAGTGACTGG 480  
QY 892 CAGCGACTCATCAACAACTATGGGGATTCGGCCCAAGAGACTCAACTTCAAACTCTTC 951  
DB 481 CAGCGACTCATCAACAACTATGGGGATTCGGCCCAAGAGACTCAACTTCAAACTCTTC 540  
QY 952 AACATCCAAGTCAAGAGAGTCAAGCAATGATGGCGTCACAACTATGCTATTAACCTT 1011  
DB 541 AACATCCAAGTCAAGAGAGTCAAGCAATGATGGCGTCACAACTATGCTATTAACCTT 600  
QY 1012 ACCAGCAGGTTCAAGTCTTCTCGGACTCGGAGTACCACTTCCGTAAGTCTCGGCTCT 1071  
DB 601 ACCAGCAGGTTCAAGTCTTCTCGGACTCGGAGTACCACTTCCGTAAGTCTCGGCTCT 660  
QY 1072 GGGCACCAGGGTGCCTCCCTCGGTTCCGGGGGAGCTGTTCAATGATTCGGCAATACGGC 1131  
DB 661 GGGCACCAGGGTGCCTCCCTCGGTTCCGGGGGAGCTGTTCAATGATTCGGCAATACGGC 720  
QY 1132 TACCTGACGCTCAACAAATGCGAGCCAGCGCTGGGAGCTTCACTCTTTTACTGCGTGGAA 1191  
DB 721 TACCTGACGCTCAACAAATGCGAGCCAGCGCTGGGAGCTTCACTCTTTTACTGCGTGGAA 780  
QY 1192 TATTTCCCTTCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTTCAGCTACACCTTTGAG 1251  
DB 781 TATTTCCCTTCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTTCAGCTACACCTTTGAG 840  
QY 1252 GAAGTCCCTTTTCCACAGCAGCTACGCGCACAGCAGAGCTTGACCGGCTGATGAATCTT 1311  
DB 841 GAAGTCCCTTTTCCACAGCAGCTACGCGCACAGCAGAGCTTGACCGGCTGATGAATCTT 900  
QY 1312 CTCATCGCAATACCTGTTATACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCCAA 1371  
DB 901 CTCATCGCAATACCTGTTATACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCCAA 960  
QY 1372 AACAGGACTTGTGTTTAGCCGTGGTCTCCAGCTGGGATGCTGTTTCAAGCCCAAAAC 1431  
DB 961 AACAGGACTTGTGTTTAGCCGTGGTCTCCAGCTGGGATGCTGTTTCAAGCCCAAAAC 1020  
QY 1432 TGCGTACTGGACCTGTTATCGGCAGCAGCGGTTTCTTAAACAAACAAACAAAC 1491  
DB 1021 TGCGTACTGGACCTGTTATCGGCAGCAGCGGTTTCTTAAACAAACAAACAAAC 1080  
QY 1492 AACAGCAATTTTACCTGGAGTGTGCTTCAAAATATAA CTTCAATGGCGTGAATCCATC 1551  
DB 1081 AACAGCAATTTTACCTGGAGTGTGCTTCAAAATATAA CTTCAATGGCGTGAATCCATC 1140  
QY 1552 ATCAACCCCTGGCACTGCTATGGCTTCACAAAGACGACGAAACAAGTTCTTTCCCATG 1611  
DB 1141 ATCAACCCCTGGCACTGCTATGGCTTCACAAAGACGACGAAACAAGTTCTTTCCCATG 1200  
QY 1612 AGCGGTGTCATGATTTTGGAAAGAGAGCGCGGAGCTTTCAAAACATGCAATGGACAAT 1671  
DB 1201 AGCGGTGTCATGATTTTGGAAAGAGAGCGCGGAGCTTTCAAAACATGCAATGGACAAT 1260  
QY 1672 GTCATGATTTACAGACGAGAGGAAA TTAAGCCACTAA CCGCTGGCCACCGAAGATTT 1731  
DB 1261 GTCATGATTTACAGACGAGAGGAAA TTAAGCCACTAA CCGCTGGCCACCGAAGATTT 1320  
QY 1732 GGGACGCTGGCACTCAATTTCCAGACGACGACGACAGCCCTCGACCGGAGATGTGCAT 1791  
DB 1321 GGGACGCTGGCACTCAATTTCCAGACGACGACGACAGCCCTCGACCGGAGATGTGCAT 1380  
QY 1792 GCTATGGGAGCAATTAACCTGGCAATGTTGGCAAGATAGACAGCTGTACCTGCAAGGTTCC 1851  
DB 1381 GCTATGGGAGCAATTAACCTGGCAATGTTGGCAAGATAGACAGCTGTACCTGCAAGGTTCC 1440  
QY 1852 ATTTGGGCCAAAATTTCTCTACACAGATGGACACTTTTCAACCCGCTCTCTTATGGGGGGC 1911  
DB 1441 ATTTGGGCCAAAATTTCTCTACACAGATGGACACTTTTCAACCCGCTCTCTTATGGGGGGC 1500

QY 1912 TTTGAGCTCAAGAACCCGCTCTCTCAGATCTCTATCAAAAAACACGCTCTTCTCGGAAT 1971  
DB 1501 TTTGAGCTCAAGAACCCGCTCTCTCAGATCTCTATCAAAAAACACGCTCTTCTCGGAAT 1560  
QY 1972 CTTCCGGCGGAGTTTTCAGCTACAAAGTTTCTTCAATCATCAACCAATCTCCACAGGA 2031  
DB 1561 CTTCCGGCGGAGTTTTCAGCTACAAAGTTTCTTCAATCATCAACCAATCTCCACAGGA 1620  
QY 2032 CAAAGTGAGTGTGGAAATTTGAATGGGAGCTGCAAGAAAGAAAAACAGCAAGCGCTGGAAATCCC 2091  
DB 1621 CAAAGTGAGTGTGGAAATTTGAATGGGAGCTGCAAGAAAGAAAAACAGCAAGCGCTGGAAATCCC 1680  
QY 2092 GAAGTGCAGTACACATCCAAATTTATGCAAAATCTGCCAAGTTGATTTTACTGTGGAAC 2151  
DB 1681 GAAGTGCAGTACACATCCAAATTTATGCAAAATCTGCCAAGTTGATTTTACTGTGGAAC 1740  
QY 2152 AATGAGCTTTTACTGAGCTCGCCCATTTGGCACCCTTACTTACCCGCTCCCTGT 2209  
DB 1741 AATGAGCTTTTACTGAGCTCGCCCATTTGGCACCCTTACTTACCCGCTCCCTGT 1798

## RESULT 6

US-09-807-802A-16  
; Sequence 16, Application US/09807802A  
; Patent No. 6759237  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,  
; FILE OF INVENTION: Vectors and Host Cells Containing Same  
; FILE REFERENCE: GNPVN.031USA  
; CURRENT APPLICATION NUMBER: US/09/807,802A  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/107,114  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: PCI/US99/25694  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 1605  
; TYPE: DNA  
; ORGANISM: AAV-1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1602)  
; OTHER INFORMATION:  
US-09-807-802A-16

Query Match 72.6%; Score 1603; DB 3; Length 1605;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 607 ATGGCTTTCAGCGGTGGCGCA CCAATGGCAGACAATAAGAGGCGCCGACGAGTGGGT 656  
DB 1 ATGGCTTTCAGCGGTGGCGCA CCAATGGCAGACAATAAGAGGCGCCGACGAGTGGGT 60  
QY 667 AATGCTCTAGGAAATTTGGCAATTCGATTCACATGGCTGGCGACAGAGTCAATCACCACC 726  
DB 61 AATGCTCTAGGAAATTTGGCAATTCGATTCACATGGCTGGCGACAGAGTCAATCACCACC 120  
QY 727 AGCACCAGCACTGGGCTTGGCCACTTAACTAACACCTCTTAAAGCAAAATCTCCAGT 786  
DB 121 AGCACCAGCACTGGGCTTGGCCACTTAACTAACACCTCTTAAAGCAAAATCTCCAGT 180  
QY 787 GCTTCAACGGGGCCAGCAACCACTACTTCTGGGTACAGCAACCCCTCGGGGTAT 846  
DB 181 GCTTCAACGGGGCCAGCAACCACTACTTCTGGGTACAGCAACCCCTCGGGGTAT 240  
QY 847 TTTGATTTCAACAGATTTCCACTGCGCACTTTTACCAGCTGACTGGCAGGACTCATCAAC 906  
DB 241 TTTGATTTCAACAGATTTCCACTGCGCACTTTTACCAGCTGACTGGCAGGACTCATCAAC 300

QY 907 AACAAATGGGGATTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAACTCAAG 966  
DB 301 AACAAATGGGGATTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAACTCAAG 360  
QY 967 GAGGTCAAGACGAATGATGGCGTCACAACTCGCTAATAAATCTTACAGACAGGTTCAA 1026  
DB 361 GAGGTCAAGACGAATGATGGCGTCACAACTCGCTAATAAATCTTACAGACAGGTTCAA 420  
QY 1027 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGCTGCTCGGCTTGGCGACAGGCTGC 1086  
DB 421 GTCTTCTCGGACTCGGAGTACAGCTTCGGTACGCTGCTCGGCTTGGCGACAGGCTGC 480  
QY 1087 CTCCTCTCGGCTTCGGCGAGCGTGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 1146  
DB 481 CTCCTCTCGGCTTCGGCGAGCGTGTTCATGATTCGCAATACGGCTACCTGACGCTCAAC 540  
QY 1147 AATGGCAGCAAGCGTGGGAGCGTTTCATCTTTTACTGCTTGGAAATATTTCCCTTCTCAG 1206  
DB 541 AATGGCAGCAAGCGTGGGAGCGTTTCATCTTTTACTGCTTGGAAATATTTCCCTTCTCAG 600  
QY 1207 ATGCTGAGAACCGGGAACAACTTTTACCTTCAGCTACACTTTGAGGAAGTGCCTTTCAC 1266  
DB 601 ATGCTGAGAACCGGGAACAACTTTTACCTTCAGCTACACTTTGAGGAAGTGCCTTTCAC 660  
QY 1267 AGCAGCTACGCGACAGCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 1326  
DB 661 AGCAGCTACGCGACAGCAGAGCCTGGACCGGCTGATGAATCCTCTCATCGACCAATAC 720  
QY 1327 CTGTATTACCTGAACAGAACTCAAACTCAAGTCCGGAAGTCCCAAAACAAAGGACTTGTG 1386  
DB 721 CTGTATTACCTGAACAGAACTCAAACTCAAGTCCGGAAGTCCCAAAACAAAGGACTTGTG 780  
QY 1387 TTTAGCGGTGCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACCTGACACC 1446  
DB 781 TTTAGCGGTGCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACCTGACACC 840  
QY 1447 TGTATTCCGACGAGCGGCTTTCAAAACAAACAGACAAACAAACAGCAATTTTACC 1506  
DB 841 TGTATTCCGACGAGCGGCTTTCAAAACAAACAGACAAACAAACAGCAATTTTACC 900  
QY 1507 TGGACTGGTGTCTCAAAATATAACCTCAATGGCGGTGAATCCATCAATCAACCTGGCACT 1566  
DB 901 TGGACTGGTGTCTCAAAATATAACCTCAATGGCGGTGAATCCATCAATCAACCTGGCACT 960  
QY 1567 GCTATGGCTTCACAAAGACGAGCAAGTGTCTTCCATGAGCGGTGTGATTT 1626  
DB 961 GCTATGGCTTCACAAAGACGAGCAAGTGTCTTCCATGAGCGGTGTGATTT 1020  
QY 1627 TTTGGAAGAGAGCGCGGAGCTTCAAACTGCAATTCGACAAATGTCTATGATTACAGAC 1686  
DB 1021 TTTGGAAGAGAGCGCGGAGCTTCAAACTGCAATTCGACAAATGTCTATGATTACAGAC 1080  
QY 1687 GAAGAGGAATTTAAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1746  
DB 1081 GAAGAGGAATTTAAAGCCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTGGCAGTC 1140  
QY 1747 AATTTCCAGACGACGACAGACCTCTGCGACCGGAGATGTGATCTATGAGGACATTA 1806  
DB 1141 AATTTCCAGACGACGACAGACCTCTGCGACCGGAGATGTGATCTATGAGGACATTA 1200  
QY 1807 CCTGGCATGGTGGCAAGATAGAGACGTGTACTGACGGGTCCCATTTTGGGCCAAATTT 1866  
DB 1201 CCTGGCATGGTGGCAAGATAGAGACGTGTACTGACGGGTCCCATTTTGGGCCAAATTT 1260  
QY 1867 CTTCAACAGATGGACACTTTTCAACCGTCTCTCTTTATGGCGGCTTTGGACTCAAGAAC 1926  
DB 1261 CTTCAACAGATGGACACTTTTCAACCGTCTCTCTTTATGGCGGCTTTGGAATCAAGAAC 1320  
QY 1927 CGGCTCTCTCAGATCTCTCAATAAAACAGCGCTGTCTCTGCGAATCTCTCGGCGGAGTTT 1986  
DB 1321 CGGCTCTCTCAGATCTCTCAATAAAACAGCGCTGTCTCTGCGAATCTCTCGGCGGAGTTT 1380  
QY 1987 TCAGCTACAAAGTTTGTCTTCAATCATCAACCACTACTCCACAGACAGATGTAGTGTGGAA 2046

DB 1381 TCAGCTACAAAGTTTGTCTTCAATCATCAACCACTACTCCACAGACAAAGTGTAGTGTGGAA 1440  
QY 2047 ATTGAATGGGAGCTGCGAGAAAGAAAACAGCAAGCGCTGGAAATCCGGAAGTGCAGTACACA 2106  
DB 1441 ATTGAATGGGAGCTGCGAGAAAGAAAACAGCAAGCGCTGGAAATCCGGAAGTGCAGTACACA 1500  
QY 2107 TCCAAATTATGCAAAATCTGCCAAGCTGTGATTTTACTGTGGACAAATGGACTTTTATACT 2166  
DB 1501 TCCAAATTATGCAAAATCTGCCAAGCTGTGATTTTACTGTGGACAAATGGACTTTTATACT 1560  
QY 2167 GAGCTCTGCCCTTATGGCACCCGTTACCTTACCCTGCCCTGT 2209  
DB 1561 GAGCTCTGCCCTTATGGCACCCGTTACCTTACCCTGCCCTGT 1603

## RESULT 7

US-09-770-315-4  
; Sequence 4, Application US/09770315  
; Patent No. 6429001  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Recombinant AAV Packaging Systems  
; FILE REFERENCE: 20263-501  
; CURRENT APPLICATION NUMBER: US/09/770,315  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,536  
; PRIOR FILING DATE: 2000-01-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 4072  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: recombinant DNA  
US-09-770-315-4

Query Match 66.0%; Score 1458.6; DB 3; Length 4072;  
Best Local Similarity 79.1%; Pred. No. 0;  
Matches 1747; Conservative 0; Mismatches 459; Indels 3; Gaps 1;

QY 1 ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60  
DB 1484 ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGAATAGA 1543  
QY 61 GAGTGTGGGACTTGAACCTGGAGCCCGGAAAGCCCAAGCCAAAGCCAAAGCAGCAAAAGCAGAC 120  
DB 1544 CAGTGTGGAGCTCAAACCTGGCCCAACCAACCAAGCCGAGAGCGGCATAAGGAC 1603  
QY 121 GACGCCCGGGTCTGGTCTTCTGGCTTCAAGTACCTCGGACCTTCAACGGACTCGAC 180  
DB 1604 GACAGCGGGTCTTGTCTTCTGGGTACAAGTACCTCGGACCTTCAACGGACTCGAC 1663  
QY 181 AAGGGGAGCGCTCAACCGCGGAGCGGAGCGGCGCTCGAGCAGCAAGGCTTACGAC 240  
DB 1664 AAGGGAGCGCGGTCAAGAGGAGCGGCGGCGCTCGAGCAGCAAGGCTTACGAC 1723  
QY 241 CAGCAGCTCAAAAGCGGGTGACAACTCGGTACCTTGGCGTATAACCAACCGCGAGCGGAGTTT 300  
DB 1724 CGGAGCTCGACAGCGGAGACAAACCGGTACCTCAAGTACAACCAACCGCGAGGTTT 1783  
QY 301 CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGGCAACCTCGGGGAGGAGCTCTTCCAG 360  
DB 1784 CAGGAGCGCTTAAAGAGATACGTCTTTTGGGGGCAACCTCGGAGGAGCTCTTCCAG 1843  
QY 361 GCGAAGAGCGGTTCTGAACTCTCGTCTGTTTGGAGGAGCGCTTAAAGCGGCTCT 420  
DB 1844 GCGAAGAGCGGTTCTTGAACCTCTGGGCTTGGTTGAGGAACCTTGAAGACGGCTCG 1903  
QY 421 GGAAGAAACGTCGCGTAGAGCAGTCGCGACAAAGAGCCAGACTCTCTCTCGGGCAGCGC 480  
DB 1904 GGAAGAAAGCGCGGTAGAGCAGCTCTCTGTGGAGCGCAGACTCTCTCTCTCGGAAACCGGA 1963





Qy		61	GAGTGTGGGACTTTGAAACCTCGAGCGCCCGAAGCCCCAACAGCAACAGCAGAAAAGCAGGAC	120
Db		2263	CAGTGTGGGAAGCTCAAACCCTGGCCCCACCAACCAGAAAGCCCAGAGCGGCATAAGGAC	2322
Qy		121	GACGGCGGGTCTGTGTCTTCCTGGCTACAAGTAACCTCGGACCTCTCAACGCACTCGAC	180
Db		2323	GACAGAGGGTCTTTGTCTCTCTGGGTACAAGTAACCTCGGACCTCTCAACGCACTCGAC	2382
Qy		181	AAGGGGAGCCCTCAACCGCGCGGACGCAAGCGGCCCTCGAGCACGAAGAAGGCTTAGCAC	240
Db		2383	AAGGAGAGCCGTCAACGAGGCGAGACGCGCGGCCCTCGAGCACGAAGAAGCTTAGCAC	2442
Qy		241	CAGCAGCTCAAGCGGGTGCAATTCGGTACCTCGGTATACCAACGCGCAGCGCGAGTTTT	300
Db		2443	CGGCGCTCGACAGCGGAGACAACCCGTATCCTCAAGTACAACCAACCGCGCGGAGTTTT	2502
Qy		301	CAGGAGGCTCGCAAGNAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
Db		2503	CAGGAGCGCTTAAAGAAGATACGTCTTTTGGGGGCAACCTCGGACGAGCAGTCTTCCAG	2562
Qy		361	GCCAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT	420
Db		2563	GCGNAAAGAGGGTCTTGAACCTCTGGGCTGGTTGAGGAACCTGTTAAGACGGTCCG	2622
Qy		421	GGAAAGAAACGTCCGTPAGAGCATGCCACAAGAGCCAGACTCTCTCGGCACTCGGC	480
Db		2623	GGNAAAGAGCGGTAGAGCATCTCTCTGTGGAGCCAGACTCTCTCGGNAACCGGA	2682
Qy		481	AAGACGGCAGAGCCCGCTAAAGAGAGACTCAATTTGGTCAGACTGGCGACTCAGAG	540
Db		2683	AAGCGGGCAGAGCCCTCGAAGAAAAGATTGAATTTGGTCAGACTGGAGACGCGAC	2742
Qy		541	TCAGTCCCGATCCACAACCTCTCGGAGAACCTCTCAGCAACCCCGCTGTGTGGGACCT	600
Db		2743	TCAGTACTGACCCCGAGCTCTCGACAGCCACGACGCGCCCTCTGTCTGGGAAC	2802
Qy		601	ACTCAATGGCTTCAGCGGTGGCGCACCAATGGCAGACAAATAAGAGGCGCGCAGCGA	660
Db		2803	AATACGATGGCTACAGGCACTGGCGCACCAATGGCAGACAAATAAGGCGCGCAGCGA	2862
Qy		661	GTGGTAAATGCTCAGGAAATTTGGCATTTGGATTCCAATGGCTGGGCGACAGAGTCATC	720
Db		2863	GTGGTAAATTTCTCGGAAATTTGGCATTTGGATTCCAATTTGGTGGGCGACAGAGTCATC	2922
Qy		721	ACCACGAGACCCGCACTCGGCTTCGCCCACTACAATAACCACTCTACAAGCAATC	780
Db		2923	ACCAACGACCCGNACTGGGCTTCGCCCACTACAATAACCACTCTACAACAAAT	2982
Qy		781	TCCAGTGTCTTCAACGGGGCGCAGCAACCACTACTTTCGGCTACAGCACCCCTCGG	840
Db		2983	TCCAGCAATCA---GGAGCTCGAACGACATCACTACTTTGGCTACAGCACCCCTTGG	3039
Qy		841	GGGTATTTTGAATTTCAACAGATTTCCAATGGCACTTTTTCACAAGTGAATGGCAGCACTC	900
Db		3040	GGGTATTTTGACTTTCAACAGATTTCCAATGGCACTTTTTCACAAGTGAATGGCAGCACTC	3099
Qy		901	ATCAACAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTTTCAACATCCA	960
Db		3100	ATCAACAACAACTGGGATTTCCGACCAAGAGACTCAACTTCAAGCTCTTTAAACATTCAA	3159
Qy		961	GTCAAGGAGGTTCAGCAAGATGTGGGCTCACAACTACGCTAATAACCTTACAGCACG	1020
Db		3160	GTCAAGGAGGTTCAGCAAGATGTGGGCTCACAGAGAGATTTGCCAATAACCTTACAGCACG	3219
Qy		1021	GTTCAGTCTTCTCGGACTCGGAGTACAGTTTCGPTACGTTCTCGGCTCTGGGCAACAG	1080
Db		3220	GTTCAGTGTTTTATGACTCGGAGTACAGTCTCCGPTACGTTCTCGGCTCTGGGCAACAG	3279
Qy		1081	GGCTGCCTCTCTCGTTTCCGGCGGACGTGTTCATGATTCGGCAATAAGGCTACCTGACG	1140
Db		3280	GGATGCTCTCCGCGTTTCCAGCAGACGTCTTATGTTGCGCAAGTATGATACCTCACC	3339
Qy		1141	CTCAACAATGGCAGCCAGCGTGGGACGTTTCATCTCTTTTACTGCCTGGAAATATTTCCCT	1200

US-09-770-315-3				US-09-770-315-3			
; Sequence 3, Application US/09770315				; Sequence 3, Application US/09770315			
; Patent No. 642901				; Patent No. 642901			
; GENERAL INFORMATION:				; GENERAL INFORMATION:			
; APPLICANT: Chiron Corporation				; APPLICANT: Chiron Corporation			
; TITLE OF INVENTION: Recombinant AAV Packaging Systems				; TITLE OF INVENTION: Recombinant AAV Packaging Systems			
; FILE REFERENCE: 20263-501				; FILE REFERENCE: 20263-501			
; CURRENT APPLICATION NUMBER: US/09/770,315				; CURRENT APPLICATION NUMBER: US/09/770,315			
; CURRENT FILING DATE: 2001-01-26				; CURRENT FILING DATE: 2001-01-26			
; PRIOR APPLICATION NUMBER: US 60/178,536				; PRIOR APPLICATION NUMBER: US 60/178,536			
; PRIOR FILING DATE: 2000-01-26				; PRIOR FILING DATE: 2000-01-26			
; NUMBER OF SEQ ID NOS: 8				; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: FastSeq for Windows Version 3.0				; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 3				; SEQ ID NO 3			
; LENGTH: 7557				; LENGTH: 7557			
; TYPE: DNA				; TYPE: DNA			
; ORGANISM: Unknown				; ORGANISM: Unknown			
; FEATURE:				; FEATURE:			
; OTHER INFORMATION: recombinant DNA				; OTHER INFORMATION: recombinant DNA			
US-09-770-315-3				US-09-770-315-3			
Query Match 66.0%; Score 1458.6; DB 3; Length 7557;				Query Match 66.0%; Score 1458.6; DB 3; Length 7557;			
Best Local Similarity 79.1%; Pred. No. 0;				Best Local Similarity 79.1%; Pred. No. 0;			
Matches 1747; Conservative 0; Mismatches 459; Indels 3; Gaps 1;				Matches 1747; Conservative 0; Mismatches 459; Indels 3; Gaps 1;			
QY	1	ATGGCTGCCGATGGTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC	60	QY	721	ACCAACAGCACCCGCACTTGGGCTTGGCCACCTACAAATAACCACTCTACAGCAAAATC	780
DB	2231	ATGGCTGCCGATGGTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC	2290	DB	2951	ACCAACAGCACCCGCACTTGGGCTTGGCCACCTACAAATAACCACTCTACAGCAAAATC	3010
QY	61	GAGTGTGGGACCTTGAACCTGGAGCCCGGAGCCCAAGCCCAACGACCAAGAGAGGAGGAC	120	QY	781	TCCAGTGTCTTCAACGGGGCCGACGACGACCACTTCTTGGGCTTACAGCACTTGGGCTTGG	840
DB	2291	CAGTGTGGGAGCTCAAACTTGGCCCAACCAACCAAGCCCGAGAGCGGCATAGGAC	2350	DB	3011	TCAGGCCAATCA---GGAGCCTCGAAACGAAATCACTACTTTGGCTTACAGCACTTGGGCTTGG	3067
QY	121	GAGCGCGGGCTCTGGTCTTCTGCTACAAGTACTCTCGGACCTTCAACGACTTCGAC	180	QY	841	GGGTATTTTGGATTTCAACAGATTCCACTGCCACTTTTACACAGCTGCTGGCAGCACTC	900
DB	2351	GACAGCAGGGGTCTTGTGCTTCTTGGGTACAAGTACTCTCGGACCTTCAACGACTTCGAC	2410	DB	3068	GGGTATTTTGGATTTCAACAGATTCCACTGCCACTTTTACACAGCTGCTGGCAGCACTC	3127
QY	181	AAGGGGAGCGCTCAACGGCGGCGAGCAGCGGCTTCAAGTACTCTCGGACCTTCAACGACTTCGAC	240	QY	901	ATCAACAACAATTTGGGGATTCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA	960
DB	2411	AAGGGGAGCGCTCAACGGGAGCAGCGGCTTCAAGTACTCTCGGACCTTCAACGACTTCGAC	2470	DB	3128	ATCAACAACAATTTGGGGATTCGGCCCAAGAGACTCAACTTTCAAACTCTTAAATTCAA	3187
QY	241	CACAGCTCAAGCGGGTGACATTCGTACTCTGGGTATACCAACGCGGACGCGGCTTCT	300	QY	961	GTCAAGGAGGTCAACGAGATGATGGCGTCACAACCATCGCTAATACCTTTACAGCAGC	1020
DB	2471	CGGAGCTCGACAGCGGAGACCAACCGTACTCTCAAGTACTCAAGTACTCTCGGACCTTCAACGACTTCGAC	2530	DB	3188	GTCAAGGAGGTCAACGAGATGATGGCGTCACAACCATCGCTAATACCTTTACAGCAGC	3247
QY	301	CAGAGGCTCTGCAAGAGATAGTCTTTTGGGGGCACTCTCGGCGAGCAGTCTTCCAG	360	QY	1021	GTTCAGAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAGTCTCTCGGCTCTGGCAGCAG	1080
DB	2531	CAGAGGCGCTTAAAGAGATAGTCTTTTGGGGGCACTCTCGGCGAGCAGTCTTCCAG	2590	DB	3248	GTTCAGAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAGTCTCTCGGCTCTGGCAGCAG	3307
QY	361	GCCAGAGCGGGTCTCGAACCTCTCGGCTCTGGGTGAGGAGCGCTAAGACGGCTCCT	420	QY	1081	GGTGGCTCTCTCGGCTTCCGGCGGAGCTTTCATGATTCGGCAATACGGCTACCTGAGC	1140
DB	2591	GCCAGAGCGGGTCTCGAACCTCTCGGCTCTGGGTGAGGAGCGCTAAGACGGCTCCT	2650	DB	3308	GGATGCTCTCGGCTTCCGAGCAGAGCTTTCATGATTCGGCAATACGGCTACCTCACC	3367
QY	421	GGAAGAAACGTCGGGTAGAGCAGTCCGCAACAGAGCCAGACTCTCTCTCGGGCATCGGC	480	QY	1141	CTCAACAATTTGGGAGCCGAGCGGTTCATCTTTTACCTTCTGCTTCTGGAATATTTCCCT	1200
DB	2651	GGAAGAAACGTCGGGTAGAGCAGTCCGCAACAGAGCCAGACTCTCTCTCGGGCATCGGC	2710	DB	3368	CTCAACAATTTGGGAGCCGAGCGGTTCATCTTTTACCTTCTGCTTCTGGAATATTTCCCT	3427
QY	481	AAGACGGCCAGAGCGCCGCTTAAAGAGACTCAATTTTGGTCTGAGCTGCGGACTCAGAG	540	QY	1201	TCTCAGATCTGAGAACGGCGCAACACTTTTACCTTTCAGCTTACACTTTTGAAGGAGTGCCT	1260
DB	2711	AAGCGGGCCAGAGCGCTTGAAGAAAGATTTGAATTTTGGTCTGAGACGCGAGAC	2770	DB	3428	TCTCAGATCTGAGAACGGCGCAACACTTTTACCTTTCAGCTTACACTTTTGAAGGAGTGCCT	3487
QY	541	TCAGTCCCGCATCCACAACTCTCGGAGAACCTTCCAGCAACCCCGCTGTGTGGGACCT	600	QY	1261	TTCCACAGAGCTACCGCGACAGCCAGAGCTTGGACCGGCTGATGATCTCTCATCGAC	1320
DB	2771	TCAGTCCCGCATCCACAACTCTCGGAGAACCTTCCAGCAACCCCGCTGTGTGGGACCT	2830	DB	3488	TTCCACAGAGCTACCGCTCAAGCAGAGCTTGGACCGCTCTCATGATCTCTCATCGAC	3547
QY	601	ACTCAATTTGGCTTTCAGCGGTGGCGCACTCAATTTGGCAGCAATTAACGAGGCGCGGACGGA	660	QY	1321	CAATACCTGATTTACCTGAAACAGAACTCAAAATCAGTCCGGAGTCCCAAAACAGGAC	1380
DB	2831	ANTACGATGCTACAGGCGAGTGGCGCACTCAATTTGGCAGCAATTAACGAGGCGCGGACGGA	2890	DB	3548	CAGTACCTGATTTACCTGAAACAGAACTCAAAATCAGTCCGGAGTCCCAAAACAGGAC	3607
QY	661	GTGGTAAATGCTCAGGAAATTTGGCATTTGGCATTTGGCATTTGGCATTTGGCATTTGGCAT	720	QY	1381	TTGCTGTTTACCGCTGGCTCTCAGCTGCGCATCTGCTTTCAGCGCCCAAAACTTGGCTACCT	1440
DB	2891	GTGGTAAATGCTCAGGAAATTTGGCATTTGGCATTTGGCATTTGGCATTTGGCATTTGGCAT	2950	DB	3608	CTTCAGTTTCTCAGCGCGGAGCGAGTGACATTCGGGACCGAGTCTAGGACTGGCTTCT	3667

Db 4028 GTTCTTCCAGGCAATGGTCTGGCAGGACAGAGATGTGTACCTTCAGGGGCCATCTGGGCA 4087  
Qy 1861 AAAATTCCTCAACAGATGGACACTTTTACCCGCTCTCTCTATATGGCGGCTTTGGACTC 1920  
Db 4088 AAGATTCCACACACGAGGACATTTTCAACCCCTCTCCCTCATGGGTGGATTGGACTT 4147  
Qy 1921 AAGAACCCGCTCTCCAGATCTCATCAAAACACGCTGTTCCTGCGAATCCTCGGGG 1980  
Db 4148 AACACCCCTCTCCACAGATCTCATCAAGAACACCCCGGTACCTGCGAATCCTTCGACC 4207  
Qy 1981 GAGTTTTCAGCTACAAAGTTTGTCTTCATTCATCACCAATACTTCACAGGACAAAGTGA 2040  
Db 4208 ACCTTCAGTGGCGCAAGTTTGTCTTCATCACACAGTACTCCACGGGACAGTCA 4267  
Qy 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGAG 2100  
Db 4268 GTGGAGATCGAGTGGGAGCTGCAGAAAGAAACAGCAACGCTGGAATCCCGAAATTCAG 4327  
Qy 2101 TACACATCCAAATATGCAAAATCTGCCAAGTGTGATTTTACTGTGGACAAACATGACTT 2160  
Db 4328 TACACTTCCAACTACAAACAGTCTGTAAATGGACTTTTACTGTGGACAACTAATGGCGTG 4387  
Qy 2161 TATACTGAGCCTCGCCCCATTGGCACCCGTTACCTTACCCGTCCCTGT 2209  
Db 4388 TATTCAGAGCCTCGCCCCATTGGCACCATACCTGACTCGTAATCTGT 4436

## RESULT 10

US-09-770-315-2  
; Sequence 2, Application US/09770315  
; Patent No. 6429001  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Recombinant AAV Packaging Systems  
; FILE REFERENCE: 20263-501  
; CURRENT APPLICATION NUMBER: US/09/770,315  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,536  
; PRIOR FILING DATE: 2000-01-26  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 8698  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: recombinant DNA  
US-09-770-315-2

Query Match 66.0%; Score 1458.6; DB 3; Length 8698;  
Best Local Similarity 79.1%; Pred. No. 0;  
Matches 1747; Conservative 0; Mismatches 459; Indels 3; Gaps 1;  
Qy 1 ATGGCTGCGCATGGTTATCTTCAGATGGCTCGAGGACAACTCTCTGAGGCACTTCGC 60  
Db 2203 ATGGCTGCGCATGGTTATCTTCAGATGGCTCGAGGACACTCTCTGAAGAAATAAG 2262  
Qy 61 GAGTGTGGGACTTGAACCTGGAGCCCGAAGCCCAAGCCAAACAGCAAAAGCAGGAC 120  
Db 2263 CAGTGTGGAAGCTCAAACTGGCCCAACCAACCAAGCCGCGAGCGGCATAGGAC 2322  
Qy 121 GACGGCGGGGTCTGGTCTCTGGGTACAAGTACTCGGACCTTCAACGGACTCGAC 180  
Db 2323 GACAGCAGGGGTCTTGTGCTTCTGGGTACAAGTACTCGGACCTTCAACGGACTCGAC 2382  
Qy 181 AAGGGGAGCCCTCAACGCGGAGCGAGCGGCTCGAGCAGCAAGCGCTACGAC 240  
Db 2383 AAGGGAGCCGCTCAACGAGGAGCAGCGCGGCTCGAGCAGCAAGCGCTACGAC 2442  
Qy 241 CAGCAGCTCAAGCGGGGTGACAACTCGGTATACCGGTATAACCAACGCGGAGTTT 300  
Db 2443 CGGCAGCTCGACAGCGGAGACAACTCGTACCTCAAGTACAAACCAACCGCGGAGTTT 2502

Qy 301 CAGGAGCGTCTGCAAGAGATACGTCCTTTTGGGGCAACCTCGGCGAGCAGTCTTCCAG 360  
Db 2503 CAGGAGCGCTTTAAGAAAGATACGTCCTTTTGGGGCAACCTCGGAGCAGCAGTCTTCCAG 2562  
Qy 361 GCCAAGAAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAAGACGGCTCCT 420  
Db 2563 GGGAAAAGAGGGTTCTTGAACCTCTGGGCTGGTTGAGGAACCTGTTTAAGACGGCTCG 2622  
Qy 421 GGAAGAAGAAAGCTCGGTAGAGAGTGGCCACAAGAGCCAGACTCTCTCTCGGCAATCGGC 480  
Db 2623 GGAAGAAAGAGCGCGGTAGAGCACTCTCTCTGGAGCCAGACTCTCTCTCGGAAACCGGA 2682  
Qy 481 AAGCAGAGCCAGCAGCGCCGCTAAAGAGACTCAATTTTGGTTCAGACTGGGAGTCAAG 540  
Db 2683 AAGCGGCGCCAGCAGCGCTGCAAGAAAAGATTTGATTTTGGTTCAGACTGGAGACGAC 2742  
Qy 541 TCAGTCCCGGATCCACAACTCTCGGAGAACTTCAGCAACCCCGCTGTGTGGGACCT 600  
Db 2743 TCAGTACCTGACCCCGAGCTCTCGGACAGCCACCGAGCCCTCTGGTCTGGAACT 2802  
Qy 601 ACTAATATGGCTTACGGCGGTGGCCACCAATGGCAGACAATTAACGAAGGCGCCGACGGA 660  
Db 2803 AATACGATGGCTTACAGGAGTGGCGCACTCAATGGCAGACAATTAACGAAGGCGCCGACGGA 2862  
Qy 661 GTGGTAAATGCTCTCAGGAAATTTGGCATTTGCGATTCACATGGCTGGCGGACAGACTC 720  
Db 2863 GTGGTAAATTTCTCGGAAATTTGGCATTTGCGATTTCCATGGATGGGCGACAGACTC 2922  
Qy 721 ACCACAGCAGCCGCACTGGGCTTTGCCACCTTACAAATAACCACTCTCAAGCAAAATC 780  
Db 2923 ACCACAGCAGCCGCACTGGGCTTTGCCACCTTACAAATAACCACTCTCAAGCAAAATC 2982  
Qy 781 TCCAGTGTTCAAACGGGGCCAGCAACGAACCACTACTTTCGGGTACAGCAACCCCTGG 840  
Db 2983 TCCAGCCCAATCA--GGAGCCTCGAAACGACAATCACTACTTTGGGTACAGCAACCCCTGG 3039  
Qy 841 GGGTATTTTGTATTCACAGATTCACATGCGCACTTTTCCACGAGTACTGGCAGGACTC 900  
Db 3040 GGGTATTTTGTATTCACAGATTCACATGCGCACTTTTCCACGAGTACTGGCAGGACTC 3099  
Qy 901 ATCAACAACAATTTGGGATTCGGGCCCCAAGAGACTCAACTTCAAACTCTTCAACATCCA 960  
Db 3100 ATCAACAACACTGGGATTCGGACCCCAAGAGACTCAACTTCAAGCTCTTTAATCAATCAA 3159  
Qy 961 GTCAAGGAGGTCAAGCAATGATGGCGTCAACCACTCGCTAATACTTACAGAGCAG 1020  
Db 3160 GTCAAGGAGGTCAAGCAATGATGGCGTCAAGCACTCGCTAATACTTACAGAGCAG 3219  
Qy 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCGGTACGCTCTCGGCTTGCAGCAG 1080  
Db 3220 GTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCGGTACGCTCTCGGCTTGCAGCAG 3279  
Qy 1081 GGCTGCCCTCCCTCCGTTCCCGGCGAGCGTTTCATGATTTCCGCAATACGGCTACCTGAG 1140  
Db 3280 GATGCTCTCCGCGCTTCCAGCAGAGCTTTCATGTTGGTCCACAGTATGATACCTCACC 3339  
Qy 1141 CTCAACAAATGGCAGCAAGCGGTGGAGCTTCACTCTTTTACTGCTTGAATATTTCCCT 1200  
Db 3340 CTGAACAAAGGAGTCAGGAGTAGGAGCTCTTCAATTTTACTGCTTGAATATTTCCCT 3399  
Qy 1201 TCTCAGATGCTGAGAAACGGGCAACCTTACCTTACGCTACCTTTCAGGAGTGCCT 1260  
Db 3400 TCTCAGATGCTGAGTACCGGAAACAACTTACCTTACGCTACCTTTCAGGAGTGCCT 3459  
Qy 1261 TTTCCAAGCAGCTAGCGCAGCAGCAGCGCTTGGACCGGCTGATGAATCTCTCATCGAC 1320  
Db 3460 TTTCCAAGCAGCTAGCGTACAGCAGCAGCTTGGACCGCTCTCATGAATCTCTCATCGAC 3519  
Qy 1321 CAATACCTGTATTAATCTGAACAGAACTCAAAATTCAGTTCGGAGTGCCTCAAAACAGGAC 1380  
Db 3520 CAGTACCTGTATTAATCTTGAGAGAAACAACTCAAACTCCAAAGTGGAAACCAACCGCAGTCAAGG 3579





```
QY 901 ATCAACAACAAATGGGGATTTCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCAA 960
Db 3030 ATCAACAACAACTGGGGATTTCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCAA 3089
QY 961 GTCAAGGAGGTTCACGACGAATGATGGCGTCACAAACCATCGCTTAATAACCTTACACGACG 1020
Db 3090 GTCAAGGAGGTTCACGACGAATGATGGCGTCACAAACCATCGCTTAATAACCTTACACGACG 3149
QY 1021 GTTCAAGTCTTTCGGACTCGAGATACAGACTTCGATGCTTCCGTCTCGGTCTCGGACGAC 1080
Db 3150 GTTCAAGTCTTTCGACTCGAGATACAGACTTCGATGCTTCCGTCTCGGTCTCGGACGAC 3209
QY 1081 GGTCTGCTCCCTTCGGTCCCGGAGCGGTTTCATGATTCGGAATACGGCTACCTGACG 1140
Db 3210 GGTCTGCTCCCGCGTTTCGAGCGAGCGTTCATGCTCCCTCAGTATGATACCTCAC 3269
QY 1141 CTCAACAATGGGACGCAAGCCGTGGGACGTTTCACTCTTTTACTGCTGGAATATTTCCCT 1200
Db 3270 CTGACAAACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCTGGAGTACTTCCCT 3329
QY 1201 TCTCAGATCTGAGAACGGGCAACAACTTTTCACTTCACTACGCTTACCTTTGAGGAAGTCT 1260
Db 3330 TCGCATGCTAAGGACTGGAAATAACTTCCAATTCAGCTATACCTTTCGAGGATGTACT 3389
QY 1261 TTCCACAGCAGCTACGCGACAGCAGAGCCTGGACCGGCTGATGAATCTCTCATCGAC 1320
Db 3390 TTTCACAGCAGCTACGCTCACAGCCAGATTTGGATCGCTTGTATGAATCTCTTATTGAT 3449
QY 1321 CAATACCTGTATTACCTGAACAGAACTCA---AAATCAGTCGGGAGTCCCAAAACAAG 1377
Db 3450 CAGTATCTGTACTCTGAACAGGCAAGGACCAACCTCTGGAACACCAACCAATCA 3509
QY 1378 GACTGCTCTTTAGCGTGGGTCTCCAGCTGGGATGTCTGTTCCAGCCCAAAACTGGCTA 1437
Db 3510 CGGCTGCTTTTTCAGCGCTGGGCTCAGTCTATGCTTTTCGAGCCAGAAATTTGGCTA 3569
QY 1438 CTGGACCTGTATTCGGCAGCAGCGGTTCTTAAACAAAACAGACACACACACACAGC 1497
Db 3570 CTGGGCGCTGTCTACGGCAACAGAGACTTTCAAGACTGCTTAAACGACCAACCAACAGT 3629
QY 1498 AATTTTACCTGGACTGGTCTTCAAAATATAACCTCAATGGGCGTGAATCCATCAAC 1557
Db 3630 AACTTTCTTGGACAGCGCCAGCAATATCATCTCAATGGCGGACTCGTGGTGAAT 3689
QY 1558 CTGGCAGCTGTATGGCTTCAACAAAGACGACGACAAAGTCTTTTCCCATGAGCGGT 1617
Db 3690 CCAGGACCACTATGGCCAGTCACAAAGGACGATGAAGAAAAATTTTCCCTATGCAACG 3749
QY 1618 GTCATGATTTTGGAAAGAGAGCGCGGAGCTTCAACACATGCTATGGACATGTCATG 1677
Db 3750 AATCTTAATTTTGGCAAGAGGGAACAACGCAAGTAAACGCAATTAGATAATGTAATG 3809
QY 1678 ATTACAGAGAGAGGAAATTAAGCCACTTAACCCCTGTGGCCACCGAAAGATTGGGACC 1737
Db 3810 ATTACGGATGAAGAGAGATTCTACACCAATCTCTGTGGCAACAGAGCATGGAAT 3869
QY 1738 GTGGCAGTCAATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCATGCTATG 1797
Db 3870 GTGGCAATAAATTCGAGAGCTCAAAATACAGCTCCACGACTGGAACGTGCAATCATCAG 3929
QY 1798 GAGGCAATTAACCTGGCATGGTGGCAAGATAGAGAGCTGTACTGCGAGGTCCTCATTTGG 1857
Db 3930 GGGGCTTACCTGGCATGGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTATCTGG 3989
QY 1858 GCCAAAATTCCTCACACAGATGACACTTTTCAACCCCTCTCTTATGGCGGCTTTGGA 1917
Db 3990 GCAAGATTTCTCACACGATGACACTTTTCACTCTCTCTCTCTGTATGGAGGCTTTGA 4049
QY 1918 CTCAGAAACCCGCTCTCTCAGATCTCTCATCAAAAAACACGCTGTTTCTGCGAATCTCCG 1977
Db 4050 CTGAAACATCCGCTCTCTCAAAATCATGATCAAAAAATACCTCCGCTAGCTGCGAATCTCCG 4109
QY 1978 GCGGAGTTTTCAGCTACAAAGTTTGCTTCAATTCATCAACCAAACTCTCCACGAGCAAGT 2037
```

```
Db 4110 ACCACCTTCAGTCGGGCAAGAGTTTGTCTTCTCATCACAGTACTCCACGGACAGTCT 4169
QY 2038 AGTGTGGAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGTGAATCCCGAAGTG 2097
Db 4170 AGCGTGGAGATCGAGTGGGAGCTGCAGAAAGAAACAGCAACCGTGAATCCCGAAAT 4229
QY 2098 CAGTACACATCCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAACATGA 2157
Db 4230 CAGTACACATTTCCAATTCACCAAGTCTGTTAATCGTGGACTTACCGTGGATTAATGCG 4289
QY 2158 CTTTATATCTAGCCTCGCCCATTTGGCACCCGTTACCTTACCTTACCGTCCCTGT 2209
Db 4290 GTGATTCAGAGCCTCGCCCATTTGGCACCCGATTTGGCACCAAGATACCTGACTCGTAATCTGT 4341

RESULT 12
US-08-254-358-1
; Sequence 1, Application US/08254358
; Patent No. 5658785
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/254,358
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5658785and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-254-358-1

Query Match 65.0%; Score 1436.4; DB 2; Length 4680;
Best Local Similarity 78.7%; Pred. No. 0;
Matches 1740; Conservative 0; Mismatches 466; Indels 4; Gaps 2;

QY 1 ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGACCAACCTCTCTGAGGGCATTCGC 60
Db 2203 ATGGCTGCGGATGGTTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAGGAATAAGA 2262
QY 61 GAGTGTGGGACTTCAACCTCGAGCCCGGAGCCCAAGCCCAACAGCAAGAGGAGGAC 120
Db 2263 CAGTGTGGAGCTCAAACTGGCCACCAACCAAGAGCCCGAGAGCGGCGATAGGAC 2322
QY 121 GACGGCCGGGGTCTGGTGTCTTCTGGCTACAGTACCTCGGACCTTCAACGGACTCGAC 180
Db 2323 GACAGCGGGGTCTTGTGCTTCTTCTGGGTACAGTACCTCGGACCTTCAACGGACTCGAC 2382
```





1 NUMBER OF SEQUENCES: 3  
2 CORRESPONDENCE ADDRESS:  
3 ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun  
4 STREET: 6300 Sears Tower, 233 S. Wacker Drive  
5 CITY: Chicago  
6 STATE: Illinois  
7 COUNTRY: USA  
8 ZIP: 60606  
9  
10 COMPUTER READABLE FORM:  
11 MEDIUM TYPE: Floppy disk  
12 COMPUTER: IBM PC compatible  
13 OPERATING SYSTEM: PC-DOS/MS-DOS  
14 SOFTWARE: PatentIn Release #1.0, Version #1.25  
15 CURRENT APPLICATION DATA:  
16 FILING DATE: 07-JUN-1995  
17 APPLICATION NUMBER: US/08/475,391  
18 CLASSIFICATION: 435  
19 PRIOR APPLICATION DATA:  
20 APPLICATION NUMBER: 08/254,358  
21 FILING DATE:  
22 ATTORNEY/AGENT INFORMATION:  
23 NAME: No. 578621land, Greta S.  
24 REGISTRATION NUMBER: 35,302  
25 REFERENCE/DOCKET NUMBER: 31975  
26 TELECOMMUNICATION INFORMATION:  
27 TELEPHONE: (312) 474-6300  
28 TELEFAX: (312) 474-0448  
29 TELEX: 25-3856  
30 INFORMATION FOR SEQ ID NO: 1:  
31 SEQUENCE CHARACTERISTICS:  
32 LENGTH: 4680 base pairs  
33 TYPE: nucleic acid  
34 STRANDEDNESS: single  
35 TOPOLOGY: linear  
36 MOLECULE TYPE: DNA (genomic)  
37 US-08-475-391-1

Query Match 65.0%; Score 1436.4; DB 2; Length 4680;  
Best Local Similarity 78.7%; Pred. No. 0;  
Matches 1740; Conservative 0; Mismatches 466; Indels 4; Gaps 2;

QY 1 ATGGCTGCGGATGGTTATCTTCAGATGGCTCGAGGACAACTCTCTGAGGCGATTGCG 60  
DB 2203 ATGGCTGCGGATGGTTATCTTCAGATGGCTCGAGGACAACTCTCTGAGGAAATAGA 2262  
QY 61 GAGTGTGGGACTTGAACCTTGAGCCCGGAAAGCCCAAGCCAAAGCCAAAGCAGGAG 120  
DB 2263 GAGTGTGGGAGCTCAACCTGGCCCAACCAACCAAGCCGCGAGCGGATAAGGAC 2322  
QY 121 GACGGCGGGGCTGGTGTCTTCTGCTACAGTACTCGGACCCCTTCAACGGACTCGAC 180  
DB 2323 GACAGAGGGGCTTGTGTCTTCTGCTTCAAGTACTCGGACCCCTTCAACGGACTCGAC 2382  
QY 181 AAGGGGAGCCGCTCAACGCGGAGCGAGCGGCGCTCGAGCAGCAGAGGCGCTACGAC 240  
DB 2383 AAGGGAGCCGCTCAACGAGGAGCAGCGCGGCGCTCGAGCAGCAGAGGCGCTACGAC 2442  
QY 241 CAGCAGCTCAAGCGGGGTGACAACTCGGTATACCTCGGTATAAACCAACCGCAGCGAGTTT 300  
DB 2443 CGGCAGCTCGACAGCGGAGACACCGGTACCTCAAGTACAAACCAACCGCAGCGAGTTT 2502  
QY 301 CAGGAGGCTCGAAGAGATAGTCTTTTGGGGGCAACCTCGGGGAGGAGTCTTCAG 360  
DB 2503 CAGGAGCGCTTAAAGAAATAGTCTTTTGGGGGCAACCTCGGAGCGAGTCTTCCAG 2562  
QY 361 GCCAAGAGCGGTTCTCGAACCTCTCGGTCTGTTGAGGAGCGCTAAGCGGCTCCT 420  
DB 2563 GCGAAAGAGGGTCTTGAACCTCTGAGGCTGTTGAGGAACTGTTAAGACGGCTCCG 2622  
QY 421 GGAAGAAAGAGTCCGGGTAGAGGAGTGGCCCAAGAGCCAGACTCCTCCTCGGCAATCGGC 480  
DB 2623 GGAAGAAAGAGCGGTAGAGGACTCTCTGTGTGAGGACAGACTCCTCCTCGGAAACCGGA 2682

QY 481 AAGACAGCCAGCAGCCGCTAAAGAGAGACTCAATTTTGGTCAGACTGGGACTCAGAG 540  
DB 2693 AAGGCGGCCAGCAGCCCTGCAAGAAAAGATTGAATTTTGGTCAGACTGGAGACGACAC 2742  
QY 541 TCAGTCCCGGATCCACAACCTCTCGGAGAACTCTCCAGCAACCCCGCTGCTGTGGGACT 600  
DB 2743 TCAGTACCTGACCCCGGCTCTCGGACAGCAACCCAGCCCTCTGCTGTGGAACT 2802  
QY 601 ACTACAATGGCTTCAAGCGGTGGCGCAACCAATGGCAGACAAATAAGAAAGGCGCGACGA 660  
DB 2803 AATAGATGGCTACAGGCACTGGCGCAACCAATGGCAGACAAATAAGAGGGGCGCGACGA 2862  
QY 661 GTGGTAATGCTCAGGAAATTTGGCAATTTGCCATTTCCATGCTGGGCGCAGAGTCAATC 720  
DB 2863 GTGGTAATTTCTCGGAAATTTGGCAATTTGCCATTTCCATGATGGGCGCAGAGTCAATC 2922  
QY 721 ACCACAGCAGCCCGCAGCTGGGCGCTTGGCCACCTTACAAATAACCAACCTCTACAAGCAATC 780  
DB 2923 ACCACAGCAGCCCGCAACCTGGGCGCTTGGCCACCTTACAAATAACCAACCTCTACAAGCAATC 2982  
QY 781 TCCAGTGTCTCAACGGGGGCGCAGCAACCAACCACTACTCTTCGGCTCAGACACCCCTGG 840  
DB 2983 TCCAGCCAAATCA--GGAGCCTCGAAACGACAAATCACTACTTTTGGCTACAGCAGCCCTGG 3039  
QY 841 GGGTATTTTGAATTTCAACAGATTTCCACTGCGCACTTTTCCACAGCTGACTGGCAGGACTC 900  
DB 3040 GGGTATTTTGAATTTCAACAGATTTCCACTGCGCACTTTTCCACAGCTGACTGGCAGGACTC 3099  
QY 901 ATCAACAACAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCAA 960  
DB 3100 ATCAACAACAATTTGGGATTTCCGACCCCAAGAGACTCAACTTTCAAGCTCTTTAAACATCAA 3159  
QY 961 GTCAAGAGGTCACGACCAATGATGGCGTCAACCAATCGCTATAAATCTTACCAAGCAG 1020  
DB 3160 GTCAAGAGGTCACGACCAATGATGGTACGACGAGATTTGCCAATAAATCTTACCAAGCAG 3219  
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCGAGTTCCTGCTAGCTCTCGGCTTGGCGACAG 1080  
DB 3220 GTTCAAGTGTTTACTGACTCGGAGTACCGAGTTCCTGCTAGCTCTCGGCTTGGCGACTCAA 3279  
QY 1081 GGTCTCCCTCCCTCCCTCCCGCGGAGCGTGTTCATGATTCGGCAATACGGCTTACCTGACG 1140  
DB 3280 GGTCTCCCTCCCGCGTTCCTCCAGCAGACGCTTTCATGTTGCCACAGTATGGATACCTCACC 3339  
QY 1141 CTCACAATGGAGCGCAAGCGGTGGAGCTTCACTCTTTTACTGCTGGAATATTTCCCT 1200  
DB 3340 CTGAACAACGGGAGTCAGGCGAGTAGGAGCTCTTCAATTTTACTGCTGAGACTTTCT 3399  
QY 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTCAGTACACCTTTTGGAGAGTGGCT 1260  
DB 3400 TCTCAGATGCTGCTACCGGAAACAACTTTTACCTTCAGTACACCTTTTGGAGAGTGGCT 3459  
QY 1261 TTCCACAGCAGCTACCGCGCACAGCCAGAGCCTTGAACCGGCTGATGAATCTCTCATCGAC 1320  
DB 3460 TTCCACAGCAGCTACCGCTCACAGCCAGAGTCTGGACCGCTCTCATGAATCTCTCATCGAC 3519  
QY 1321 CAATACCTGTATTAACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGGCCAAACAGGAC 1380  
DB 3520 CAGTACCTGTATTAACCTGAGCAGAAACAACTTCAAGTGGAAACCAACCAACGAGTCAAGG 3579  
QY 1381 TTGCTGTTTAGCGGTGGGTCTCCAGCTGGCAATGCTGTTCAGCCCAAACTCTGGCTTACCT 1440  
DB 3580 CTTTCAGTTTTTTCAGCGCGGAGCGAGTGAATTCGGGACCACTCTAGGAACTTGGCTTCT 3639  
QY 1441 GGACCTCTGTTATTCGGCAGCAGCGGCTTTTCAAAAACAAAAACAGACAAACAAACAGCAAT 1500  
DB 3640 GGACCTCTGTACCGCCAGCAGCGAGTATCAAGAGCATCTCGGAAATAACAACAGAGTAA 3699  
QY 1501 TTTTACTGAGCTGGTCTTCAAAATATAACCTCAATGGGCGGTGAATCTCATCAACCT 1560  
DB 3700 TACTCTGAGCTGGAGTACCAAGTACCACTCAATGGCAGAGACTCTCTGGTGAATCCG 3759  
QY 1561 GGCACTGCTATGCGCTCACACAAAGCAGCAAGCAAGTCTTCTTCCCATGAGCGGTGC 1620

Db 3760 GGGCCGCCATGGCAAGCCACAGGACGATGAAGAAAGTTTTTCTCCAGAGCGGGTT 3819  
Qy 1621 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAACACATGCTTGGACATGTCATGATT 1680  
Db 3820 CTCATCTTTGGAGCAAGCTCAGAGAAACAAATGTGAACATTTGAAAAGTCTGATT 3879  
Qy 1681 ACAGACGAAGAGAAATTAAGCCACTAAACCTCTGTGGCCACCGAAAGATTTGGACCGTG 1740  
Db 3880 ACAGACGAAGAGAAATCGAAACAAATCCCGTGGCTACGAGCAGTATGTTCTGTA 3939  
Qy 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACAGACCCCTGGCCACCGGAGATGTGCTATGGGA 1800  
Db 3940 TCTACCAACCTCCAGAGAGGCAACAGACAGCAGCTACCGCAGATGTCAACACACAAGGC 3999  
Qy 1801 GCATTACCTGGCATGTGTGGCAAGATAGAGAGCTGTACTGTGAGGTCCTCATTTGGGCC 1860  
Db 4000 GTTCTCCAGGATGTGTGGCAGGACAGATGTGTACTTTCAGGGGGCCATCTGGGA 4059  
Qy 1861 AAAATTCCTCACACAGATGGACATTTTCAACCCGTCTCTCTTTATGGCGGCTTTGGACTC 1920  
Db 4060 AGATTTCCACACACGACGAGACATTTTCAACCCCTCTCCCTCATGGGTGGATTCGGACTT 4119  
Qy 1921 AGAACCCGCTCTCAGATCTCTATCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1980  
Db 4120 AAACACCTCTCTCCACAGATTTCTCATCAAGAAACACACCCCGGTACTCTGGAATCTCTGACC 4179  
Qy 1981 GAGTTTTCAGCTACAAGTTGTCTTCTATCATCAACCAACCAACCAACCAACCAACCAAC 2039  
Db 4180 ACCTTCAGTGGCGCAAGTTGTCTCTTCTCATCACACAGTACTTCCACGGGACACGGTCAG 4239  
Qy 2040 TGTGGAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAACTGCA 2099  
Db 4240 CGTGGAGATCGAGTGGGAGCTGCAGAAAGAAACAGCAAGCGTGGNATCCCGAATTC 4299  
Qy 2100 GTACACATCAATATGCAAAATCTGCCAAGTTGATTTTACGTGGACAAACATGGACT 2159  
Db 4300 GTACACTTCAACTACAACAAGTCTGTAACTGTGGACTTACCGTGGATACTAATGGCGT 4359  
Qy 2160 TTATCTAGCTCGCCCATTTGGCACCGCTTACCTTACCTTACCGTCCCTGT 2209  
Db 4360 GTATTACAGGCCTCGCCCAATGGCACCATGACCTGACTCGTAATCTGT 4409

RESULT 14  
US-08-709-609-1  
; Sequence 1, Application US/08709609  
; Patent No. 5858775  
; GENERAL INFORMATION:  
; APPLICANT: Johnson, Philip R.  
; TITLE OF INVENTION: Adeno-Associated Virus Materials and  
; TITLE OF INVENTION: Methods  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 S. Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/709,609  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5858775and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 31975

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 474-6300  
; TELEFAX: (312) 474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4680 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-709-609-1  
Query Match 65.0%; Score 1436.4; DB 2; Length 4680;  
Best Local Similarity 78.7%; Pred. No. 0;  
Matches 1740; Conservative 0; Mismatches 466; Indels 4; Gaps 2;  
Qy 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACCAACCTCTCTGAGGCGATTCCG 60  
Db 2203 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACACATCTCTCTGAGGAATAGA 2262  
Qy 61 GAGTGTGGGACTTGAACCTGGAGCCCCGAAAGCCCAAGCCAAACAGCAAAAGCAGGAC 120  
Db 2263 CAGTGTGGAAGCTCAACCTGGCCACCACCACCAAGCCCGCAGAGCGGATAGGAC 2322  
Qy 121 GACGCGCGGGTCTGGTGTCTTCTGGCTACAAAGTACCTCGGACCTTCAACGGAATCGAC 180  
Db 2323 GACAGCAGGGTCTTGTGCTTCTGGGTACAAAGTACCTCGGACCTTCAACGGAATCGAC 2382  
Qy 181 AAGGGGAGCCCCGTCAACGCGGGGAGCGAGCGGCCCTCGACACGACAGGCGCTACGAC 240  
Db 2383 AAGGGAGAGCCGGTCAACGAGCGACGCGCGGCCCTCGACACGACAGCAAAAGCTACGAC 2442  
Qy 241 CAGCAGCTCAAAAGCGGGTGACAAATCCGTACCTCGGTATAACACACGCCGACGCCAGTTT 300  
Db 2443 CGCAGCTCGACAGCGGAGACAAACCGGTACCTCAAGTACAAACACGCCGACGCCGAGTTT 2502  
Qy 301 CAGGAGCGTCTCAAGAAGATACGTCTTTTGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360  
Db 2503 CAGGAGCGCTTAAAGAGATACGTCTTTTGGGGCAACCTCGGACGAGCAGTCTTCCAG 2562  
Qy 361 GCCAAGAGCGGGTCTCGAACTCTCGTCTGGTTGAGAGGCGCTAAGCGGCTCCT 420  
Db 2563 GCGAAGAGAGGGTCTTGAACCTCTGGGCCCTGGTTGAGGAACCTGTTAAGCGGCTCG 2622  
Qy 421 GGAAGAGAAAGTCCGGTAGAGCAGTCGCCAACAAGACCCAGACTCTCTCGGCAATCGG 480  
Db 2623 GGAAGAGAGCGCGGTAGAGCAGTCTCTGTGGAGCCAGACTCTCTCTCGGAAACCGGA 2682  
Qy 481 AAGACAGGCGCAGCGCCCTCAAAAAGAGACTCAATTTTGGTACAGCTGGCGACTCAGAG 540  
Db 2683 AAGCGGGCCAGCAGCCTTCAAGAAAAGATTTGAATTTTGGTCAGACTGGAGACGAGAC 2742  
Qy 541 TCAGTCCCCGATCCACAACTCTCGAGAAACCTTCGAGAAACCCCGCTCTGTGGGACCT 600  
Db 2743 TCAGTACCTGACCCCCCAGCCTCTCGGACACCCACAGAGCCCCCTCTGTGTGGAACT 2802  
Qy 601 ACTACAATGCTTCAGGCGGTGGCGCACCAATAGGACAGCAATAACGAAGCGCGCAGCGA 660  
Db 2803 AATACGATGGCTACAGGCAAGTGGCGCAACCAATGGCAGACANAAACGAGGGCGCGACGGA 2862  
Qy 661 GTGGGTAATCCCTCAGGAAATTTGGCAATTCGATTTCCATGCTGGCGGACAGAGTCATC 720  
Db 2863 GTGGGTAATTCCTCCGGAATTTGGCATTGGCATTCGATTTCCATGATGGCGACAGTCTC 2922  
Qy 721 ACCACGACCGCGACCTTGGGCTTGGCCACCTACATAACCACTCTTACAGGAATC 780  
Db 2923 ACCACGACCGCAACCTTGGGCTTGGCCACCTACATAACCACTCTTACAAACAAAT 2982  
Qy 781 TCCAGTGTCTCAACGGGGGCGCAGCAGACCACTACTTCTGGCTACAGACCCCTCG 840  
Db 2983 TCCAGCCAATCA---GGAGCTCTGAAACGACAAATCACTACTTTGGCTACAGACCCCTTG 3039



QY 121 GACGGCCGGGCTGTGTTCTTCTGCTACAGTACCTCGGACCCCTTCAACGAGCTCGAC 180  
DB 2323 GACAGCAGGGGCTTGTGTTCTTCTGGGTACAGTACCTCGGACCCCTTCAACGAGCTCGAC 2382  
QY 181 AAGGGGAGCCGCTCAACGGGGGAGCGAGCGGCCCTCGAGCACGACGAGCCCTACGAC 240  
DB 2383 AAGGGGAGCCGCTCAACGAGGAGAGCGCGCGGCCCTCGAGCACGACGAGCCCTACGAC 2442  
QY 241 CAGCAGCTCAAGCGGGTGACAACTCCGTACCTCGGTATACCAACGCGGACGCGGAGTTT 300  
DB 2443 CGGCAGCTCGACAGCGGAGACACCCGTACCTCAAGTACACACGACGCGGAGTTT 2502  
QY 301 CAGGAGCGCTGCAAGAAATACGTTCTTTTGGGGCAACTCGGGGAGAGTCTTCCAG 360  
DB 2503 CAGGAGCGCTTAAAGAAATACGTTCTTTTGGGGCAACTCGGAGGAGAGTCTTCCAG 2562  
QY 361 GCCAAGAGGGGTTCTCGAAGCTCTCGTCTGTTTGGAGAGGCGCTAAGCGGCTCCT 420  
DB 2563 GCGAAAAGAGGGTTCTTGAACCTCTGGGCTTGGTGAAGAACCTGTTAAGAGGGCTCGG 2622  
QY 421 GGAAGAAAGCTCCGGTAGAGCAGTCGCCAACAGAGCCAGACTCTCTCGGGCATCGGC 480  
DB 2623 GGAAAAAGAGGCGGTAGAGCACTCTCTGTGGAGCCAGACTCTCTCGGAAACGGA 2682  
QY 481 AAGACAGGCGAGCGCCCGCTAAAAAGAGACTCAATTTTGGTGCAGACTGCGGACTCAGAG 540  
DB 2683 AAGGCGGCGAGCAGCTTGCAGAAAAGATTGAATTTTGGTGCAGACTGAGAGCGCAGAC 2742  
QY 541 TCAGTCCCCGATCCACACTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGACCT 600  
DB 2743 TCAGTACCTGACCCCGAGCTCTCGAGACGCCACCGAGCGCCCTCTGCTGTGGAACT 2802  
QY 601 ACTACAAATGCTTCAGCGGTGCGCACCAATGGCAGACAAATAAGAGGCGCGGACGGA 660  
DB 2803 AATACGATGCTACAGCAGTGGCGCACCAATGGCAGACAAATAAGAGGCGCGGACGGA 2862  
QY 661 GTGGGTAATGCTCAGGAAATGGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGC 720  
DB 2863 GTGGGTAATTTCTCCGAAATGGCAATGGCAATGGCAATGGCAATGGCAATGGCAATGC 2922  
QY 721 ACCACAGCACCCGACCTGGGCTTGCACCACTACAAATACCACTCTACAGCAATC 780  
DB 2923 ACCACAGCACCCGACCTGGGCTTGCACCACTACAAATACCACTCTACAAACAAAT 2982  
QY 781 TCAGTGTCTTACAGGGGCGCAGACGAAACCACTACTTTCGGCTACAGCACCCCTGG 840  
DB 2983 TCCAGGCAATCA--GGAGGCTCGAAGCAATCACTACTTTCGGCTACAGCACCCCTTGG 3039  
QY 841 GGGTATTTTGTATTTCAACAGATTCACCTGCGCACTTTTTCACAGCTGAGTGGCAGCACTC 900  
DB 3040 GGGTATTTTGTATTTCAACAGATTTCCACTGCGCACTTTTTCACAGCTGAGTGGCAGCACTC 3099  
QY 901 ATCAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960  
DB 3100 ATCAACAACTTGGGATTTCCGACCCCAAGAGACTCAACTTCAAGCTCTTTAAATTCMA 3159  
QY 961 GTCAAGGAGTCAAGCAATGATGGCTCACAACCACTCCGTAATACCTTACCAGCAG 1020  
DB 3160 GTCAAGGAGTCAAGCAATGATGGCTCACAACCACTCCGTAATACCTTACCAGCAG 3219  
QY 1021 GTTCAAGTCTTTCGAGCTCGGAGTACAGCTTCCGTAAGTCTTCGCTCGGCTCGGCAAC 1080  
DB 3220 GTTCAAGTCTTTCGAGCTCGGAGTACAGCTTCCGTAAGTCTTCGCTCGGCTCGGCAAC 3279  
QY 1081 GGTGCTCTCCCTCGGTTCCGGGAGAGTGTTCAGATTCGGAATACGGCTACCTGAGC 1140  
DB 3280 GGTGCTCTCCCTCGGTTCCGGGAGAGTGTTCAGATTCGGAATACGGCTACCTGAGC 3339  
QY 1141 CTCAACAAATGGCAGCGGCTGGGAGGTTTCATCTTTTACTGCTGGAAATATTTCCCT 1200  
DB 3340 CTGAACAAAGGAGTCAAGGAGTAGGACGCTTTCATTTTACTGCTGGAGTACTTTCT 3399  
QY 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTTCAGCTACACCTTTGAGGAAAGTGCCT 1260

DB 3400 TCTCAGATGCTCGTACCGGAACAACTTTTACCTTACGCTACACTTTTGGAGCGTTCT 3459  
QY 1261 TTCCACAGCAGCTACCGGCACACGAGCCCTGGAGCCGCTGATGAATCCTCTCATCGAC 1320  
DB 3460 TTCCACAGCAGCTACCGGCACACGAGCCGCTGGAGCCGCTGATGAATCCTCTCATCGAC 3519  
QY 1321 CAATACCTGTATTATTCCTGAACAGAACTCAAAATCAGTCCGGAGTGGCCCAAAACAGGAC 1380  
DB 3520 CAGTACCTGTATTATTCCTGAACAGAACTCAAAATCAGTCCGGAGTGGCCCAAAACAGGAC 3579  
QY 1381 TTGCTGTTTACCGGTGGGTCTCCAGCTGGCATGTCTGTTACGCCCAAACTGGCTACCT 1440  
DB 3580 CTTCAAGTTTCTCAGCGCGGAGCGAGTGACATTCGGGACCACTTAGGAACCTGGCTTCT 3639  
QY 1441 GGACCTGTATTATTCGGCAGCAGCGGCTTTCTAAAAAAGAGAAACAGAAACAGCAAT 1500  
DB 3640 GGACCTGTATTATTCGGCAGCAGCGGAGTATCAAGAGCATCTCGCGATACAAACAGTGA 3699  
QY 1501 TTTCCTGAGCTGGGTCTTCAAAATATACCTCAATGGGCGTGAATCCATCATCAACCT 1560  
DB 3700 TACTCGTGGACTGGAGCTACCAAGTACCACTCAATGGCAGAGACTCTCTGGTGAATCCG 3759  
QY 1561 GGCACCTGTATGGCTTCACAAAAGCAGAGAGCAAGTCTTTCCTCCATGAGCGGTGC 1620  
DB 3760 GGGCCGCGCATGGCAAGCCACAGGACGATGAAGAAAGTCTTTCCTCAGAGCGGGTT 3819  
QY 1621 ATGATTTTGGAAAAGAGAGCGCCGAGCTTCAAAACACTTGCATTTGGCAATGTCTATGAT 1680  
DB 3820 CTCATCTTTGGAGAGCAAGCTCAGAGAAACAAATGTGAAATTTGAAAAGGTCTATGAT 3879  
QY 1681 ACAGACGAAGAGAAATTAAGCCACTAACTGTGGCCACCGAAAGATTTGGGACCGTG 1740  
DB 3880 ACAGACGAAGAGAAATCGGAACAACTCCGCTGCTACGAGCAGTATGTTCTGTGA 3939  
QY 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACCTTCGACCGGAGAGTGTGATGCTATGGGA 1800  
DB 3940 TCTACCAACTTCCAGAGAGGCAACAGACAGCTTACCGCAGATGTCAACACACAAGGC 3999  
QY 1801 GCATTTACCTGGCATGTGGCAAGATAGAGACGTGTACTTCAGGCGTCCCAATTTGGGCG 1860  
DB 4000 GTTCTTCAGGCAATGTGTGGCAGCAGAGATGTGTACTTTCAGGCGCCCATCTGGGCA 4059  
QY 1861 AAAATTCCTCACACAGATGACATTTTCAACCGCTCTCTCTTATGGCGGCTTTGACTC 1920  
DB 4060 AAGATTTCCACACAGCAGCAGCAATTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 4119  
QY 1921 AAGAACCCGCTCTCAGATCTCATCAAAACACAGCCTGTTCCTCGGAATCTCTCGGCG 1980  
DB 4120 AAACACCTCTCTCAGATCTCATCAAGAAACACCCCGGTACCTCGGAATCTCTCGACC 4179  
QY 1981 GAGTTTTCAGCTACAAAGTTTGTCTTCACTACCCCAATACCTCCACAGGACA-AGTGAG 2039  
DB 4180 ACCTTCAGTGGCGCAAGTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4239  
QY 2040 TGTGGAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGA 2099  
DB 4240 CGTGGAGATCGAGTGGGAGCTGCAGAAAGAAACAGCAACCGTGAATCCCGAATTC 4299  
QY 2100 GTACACATCAATATGCAAAATCTGCCAAGTGTGATTTTACTGTGGAACAACTGACT 2159  
DB 4300 GTACACTTCAACTACAAAGTCTGTTAATCGTGGACTTACCGTGGATCTAATGGCGT 4359  
QY 2160 TTATACTGAGCTCGCCCGCATTTGGCACCGTTTACCTTACCCGTCCTCTGT 2209  
DB 4360 GTATTTCAGAGCTCGCCCGCATTTGGCACAGATACCTGACTCTGTAATCTGT 4409

Search completed: November 29, 2005, 14:27:45

Job time : 320.186 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 08:46:14 ; Search time 1033.63 Seconds  
(without alignments)  
14243.261 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_2223\_4431

Perfect score: 2209  
Sequence: 1 atggctgcgaagggttatct.....ttaccttaccgctccctgt 2209

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_21.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*
- 14: Geneseqn2005as.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2209	100.0	2211	3	AAD00777
2	2209	100.0	2211	14	Adz27052 Adeno-ass
3	2209	100.0	2211	14	Adz26929 Adeno-ass
4	2209	100.0	4347	13	Adw39398 Adeno-ass
5	2209	100.0	4347	14	Adz46594 HSV-AAV s
6	2209	100.0	4718	3	AAD00772
7	2209	100.0	4718	10	Adz76507 Adeno-ass
8	2209	100.0	4718	10	Adl13984 Adeno-ass
9	2209	100.0	4718	12	Adg39758 AAV-1 gen
10	2202.6	99.7	7447	8	Acc58477 Vector pF
11	2196.2	99.4	2211	14	Adz26928 Adeno-ass
12	2124	96.2	2214	14	Adz26930 Adeno-ass
13	2109.8	95.5	2211	14	Adz26931 Adeno-ass
14	2105	95.3	2211	14	Adz27053 Adeno-ass
15	2105	95.3	4239	13	Adw39402 Adeno-ass
16	2105	95.3	4239	14	Adz46598 HSV-AAV s
17	2105	95.3	4683	4	Aaf23749 AAV6 DNA
18	2105	95.3	4683	10	Adl13983 Adeno-ass
19	2105	95.3	4683	12	Adg39763 AAV-6 gen

20	2100.2	95.1	2211	14	Adz26932
21	1798	81.4	1800	3	AAD00778
22	1639	74.2	2214	14	Adz26890
23	1620	73.3	2214	14	Adz26891
24	1616.8	73.2	3122	10	Adz76546
25	1616.6	73.2	2214	14	Adz26887
26	1615.2	73.1	3117	10	Adz76544
27	1615.2	73.1	3122	10	Adz76543
28	1613.6	73.0	3121	10	Adz76545
29	1611.8	73.0	2214	14	Adz27063
30	1611.8	73.0	4721	10	Adz76502
31	1611.8	73.0	4721	12	Adg39764
32	1611.8	73.0	4721	14	Adz76509
33	1611.8	73.0	4721	14	Adz27030
34	1607	72.7	2214	14	Adz26885
35	1607	72.7	2217	14	Adz26869
36	1605.4	72.7	2214	14	Adz26886
37	1605.4	72.7	2217	14	Adz26892
38	1605.4	72.7	3127	10	Adz76516
39	1603.8	72.6	2214	14	Adz26882
40	1603.8	72.6	2217	14	Adz26875
41	1603	72.6	1605	3	AAD00779
42	1602.2	72.5	2214	14	Adz26889
43	1602.2	72.5	2217	14	Adz26865
44	1602.2	72.5	2217	14	Adz26866
45	1600.6	72.5	2217	14	Adz26867

ALIGNMENTS

RESULT 1

AAD00777

ID AAD00777 standard; DNA; 2211 BP.

XX

AC AAD00777;

XX

DT 08-SEP-2000 (first entry)

XX

DE Adeno-associated virus serotype 1 capsid protein VP1 DNA.

XX

KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein; cap protein; recombinant viral vector; gene delivery; gene therapy;

KW vaccine; transgene; VP1; ss.

XX

OS Adeno associated virus serotype 1.

XX

FH Key Location/Qualifiers

FT CDS

FT 1..2211

FT /\*tag= a

FT /product= "VP1 protein"

XX

PN WO200028061-A2.

XX

PD 18-MAY-2000.

XX

PF 02-NOV-1999; 99WO-US025694.

XX

PR 05-NOV-1998; 98US-0107114P.

XX

PA (UYPE-) UNIV PENNSYLVANIA.

XX

PI Wilson JM, Xiao W,

XX

DR WPI; 2000-376571/32.

DR P-PSDB; AAY71167.

XX

PT Novel adeno-associated virus serotype 1 polynucleotide useful for preparation of medicament for delivery of a transgene to a host.

XX

PS Claim 10; Page 83-87; 108pp; English.

XX

CC The patent discloses an adeno-associated virus serotype 1 (AAV-1) DNA



CC which is characterised by two inverted terminal repeats (ITR) and open  
CC reading frames for rep and capsid (cap) proteins. The rep reading frame  
CC encodes four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap  
CC reading frame encodes three structural proteins, VP1, VP2 and VP3. The  
CC AAV-1 sequence or its fragments particularly ITRs, rep and cap coding  
CC regions, are useful in production of recombinant viral vectors for gene  
CC delivery. These vectors can be used as gene therapy vectors, vaccine  
CC vectors or antisense delivery vectors. The AAV-1 does not induce the  
CC formation of neutralising antibodies specific to any serotype of AAV  
CC hence is useful for transducing host cells, and in the preparation of a  
CC medicament for the delivery of transgene to a host. The present sequence  
CC is an AAV-1 DNA encoding a cap protein VP1 which is useful in the  
CC production of recombinant viral vector for gene delivery  
XX

8Q Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 U; 0 Other;

Query Match		100.0%	Score 2209;	DB 3;	Length 2211;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 2209;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCAATCGC	60		
DB	1	ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCAATCGC	60		
QY	61	GAGTGTGGGACTTGAAACCTGGAGCCCGAAGCCCAAGCCCAAGCCCAAGCAAGCAGGAC	120		
DB	61	GAGTGTGGGACTTGAAACCTGGAGCCCGAAGCCCGAAGCCCAAGCCCAAGCAAGCAGGAC	120		
QY	121	GACGGCCGGGGTCTGGTGTCTCTGGCTACAAGTACTCTGGAGCCCTTCAACGGACTCGAC	180		
DB	121	GACGGCCGGGGTCTGGTGTCTCTGGCTACAAGTACTCTGGAGCCCTTCAACGGACTCGAC	180		
QY	181	AAGGGGAGCCGCTCAACCGGGCGAGCGAGCGGCCCTCGAGCAGCAAGGCGCTACGAC	240		
DB	181	AAGGGGAGCCGCTCAACCGGGCGAGCGAGCGGCCCTCGAGCAGCAAGGCGCTACGAC	240		
QY	241	CAGCAGCTCAAGCGGGTGACATCCGTACCTCGCGGTATAACACCGCCGACGCGAGTTT	300		
DB	241	CAGCAGCTCAAGCGGGTGACATCCGTACCTCGCGGTATAACACCGCCGACGCGAGTTT	300		
QY	301	CAGGAGCGTCTGCAAGAAGATAGCTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360		
DB	301	CAGGAGCGTCTGCAAGAAGATAGCTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360		
QY	361	GCCAGAAGCGGGTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT	420		
DB	361	GCCAGAAGCGGGTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT	420		
QY	421	GGAAAGAAAGCTCCGGTAGAGCAGTCGCCAAGAGCCAGACTCTCTCCCTCGGSCATCGGC	480		
DB	421	GGAAAGAAAGCTCCGGTAGAGCAGTCGCCAAGAGCCAGACTCTCTCTCCCTCGGSCATCGGC	480		
QY	481	AAGACAGGCCAGCAGCCCGCTAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG	540		
DB	481	AAGACAGGCCAGCAGCCCGCTAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG	540		
QY	541	TGAGTCCCGATCCCAAACTCTCGGAGAACTCTCAGCAACCCCGCTGTGTGGGACCT	600		
DB	541	TGAGTCCCGATCCCAAACTCTCGGAGAACTCTCAGCAACCCCGCTGTGTGGGACCT	600		
QY	601	ACTACAAATGGCTTCAGCGGTGGCGCAACCAATGGCAGACAATAACGAAGCGCCGACGGA	660		
DB	601	ACTACAAATGGCTTCAGCGGTGGCGCAACCAATGGCAGACAATAACGAAGCGCCGACGGA	660		
QY	661	GTGGGTAAATGCCCTCAGGAAATTTGGCAATTCGATTTCCACATGGCTGGCGGACAGATCATC	720		
DB	661	GTGGGTAAATGCCCTCAGGAAATTTGGCAATTTCCACATGGCTGGCGGACAGATCATC	720		
QY	721	ACCAACAGACCCCGACCTTGGGCTTGGCCCACTCAATAAACCACTCTTACAGAAATC	780		
DB	721	ACCAACAGACCCCGACCTTGGGCTTGGCCCACTCAATAAACCACTCTTACAGAAATC	780		
QY	781	TCCAGTGTCTTCAACGGGGGCGACGACCAACCACTACTCTGGCTACAGCAACCCCTCG	840		

781	Db	 TCCAGTGTCTTCAACGGGGGCGACGAA CGA CAACCACTACTTTCGGCTTACGACACCCCTGG	840
841	Qy	 GGGTATTTTGTATTTCAACAGATTTCCACTGCGCACTTTTACCACGCTGACATGGCAGGCACTC	900
841	Db	 GGGTATTTTGTATTTCAACAGATTTCCACTGCCACTTTTACCACGCTGACTGGCAGGCACTC	900
901	Qy	 ATCAACAACAAATTTGGGGATTTCCGGCCCAAGAGACTCAAACTTTCAAACCTTTCAACATCCAA	960
901	Db	 ATCAACAACAAATTTGGGGATTTCCGGCCCAAGAGACTCAAACTTTCAAACCTTTCAACATCCAA	960
961	Qy	 GTCAAGGAGGTCAAGACGAATGATGGCGTCACAACCATGCTTAATTAACCTTACCAGGACG	1020
961	Db	 GTCAAGGAGGTCAAGACGAATGATGGCGTCACAACCATGCTTAATTAACCTTACCAGGACG	1020
1021	Qy	 GTTTCAAGTCTTCTTCGGACTCGGAGTACCAAGCTTCGGTTCGTCCTCGGCTCTCGGCACCG	1080
1021	Db	 GTTTCAAGTCTTCTTCGGACTCGGAGTACCAAGCTTCGGTTCGTCCTCGGCTCTCGGCACCG	1080
1081	Qy	 GGCTGCTCCCTCCGTTCCGGCGGACGTTTTCATGATTTCCGCAATACGGCTACCTGAGC	1140
1081	Db	 GGCTGCTCCCTCCGTTCCGGCGGACGTTTTCATGATTTCCGCAATACGGCTACCTGAGC	1140
1141	Qy	 CTCAACAATGGCAGCAAGCCGCTGGGACGTTTCATGCTTTTATCTGCTTGGAAATATTTCCT	1200
1141	Db	 CTCAACAATGGCAGCAAGCCGCTGGGACGTTTCATGCTTTTATCTGCTTGGAAATATTTCCT	1200
1201	Qy	 TCTCAGATGCTGAGACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT	1260
1201	Db	 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT	1260
1261	Qy	 TTCCAAGCAGCTTACCGGCAAGCCGAGAGCTTCGGACCGGCTGATGAATCTCTCATCGAC	1320
1261	Db	 TTCCAAGCAGCTTACCGGCAAGCCGAGAGCTTCGGACCGGCTGATGAATCTCTCATCGAC	1320
1321	Qy	 CAATACTGTGTTATCTGAAACAGAACTCAAAATCAATCCGGAAGTGCACAAAACAAGGAC	1380
1321	Db	 CAATACTGTGTTATCTGAAACAGAACTCAAAATCAATCCGGAAGTGCACAAAACAAGGAC	1380
1381	Qy	 TTGCTGTTTACCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCGCCAAAACCTGGCTACCT	1440
1381	Db	 TTGCTGTTTACCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCGCCAAAACCTGGCTACCT	1440
1441	Qy	 GGACCTGTGTTATCGGCGAGCGCGTTCCTAAAAAAGAGACAGCAACAAAGAGCAAT	1500
1441	Db	 GGACCTGTGTTATCGGCGAGCGCGTTCCTAAAAAAGAGACAGCAACAAAGAGCAAT	1500
1501	Qy	 TTTACTCTGGACTGGTGTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCT	1560
1501	Db	 TTTACTCTGGACTGGTGTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCT	1560
1561	Qy	 GGCACTGCTATGGCTCTACAAAGACGAGAGCAAGTTCCTTTCCCATGAGCGGTGTC	1620
1561	Db	 GGCACTGCTATGGCTCTACAAAGACGAGAGCAAGTTCCTTTCCCATGAGCGGTGTC	1620
1621	Qy	 ATGATTTTGGAAAAAGAGCGCGGAGCTTCAAACTGCAATTTGGACAATGTGATGATTT	1680
1621	Db	 ATGATTTTGGAAAAAGAGCGCGGAGCTTCAAACTGCAATTTGGACAATGTGATGATTT	1680
1681	Qy	 ACAGACGAAGAGGAAATTAAGCCCTGTGGCCACCGAAGATTTGGGACCGTG	1740
1681	Db	 ACAGACGAAGAGGAAATTAAGCCCTGTGGCCACCGAAGATTTGGGACCGTG	1740
1741	Qy	 GCAGTCAATTTCCAGAGCAGCAGACCTTCGCGACCGGAGATGTGCAATGCTATGGGA	1800
1741	Db	 GCAGTCAATTTCCAGAGCAGCAGACCTTCGCGACCGGAGATGTGCAATGCTATGGGA	1800
1801	Qy	 GCATTTACCTGGCATGGTGGCAAGATAGAGACGTTTACCTCGAGGGTCCCATTTGGGCC	1860
1801	Db	 GCATTTACCTGGCATGGTGGCAAGATAGAGACGTTTACCTCGAGGGTCCCATTTGGGCC	1860
1861	Qy	 AAAAATTCCTCAACAGATGGACACTTTTCAACCCGCTCTCTCTTATGGCGGCTTTGGACTC	1920

Db 1861 AAAATTCTCACACAGATGGACACTTTTCACCCGCTCTCTCTATTATGGCGGCTTTGGACTC 1920  
QY 1921 AAGAACCCGCTCTCTCAGATCTCTATCAAAAACACGCGCTGTCTCTCGAATCTCTCGGG 1980  
Db 1921 AAGAACCCGCTCTCTCAGATCTCTATCAAAAACACGCGCTGTCTCTCGAATCTCTCGGG 1980  
QY 1981 GAGTTTTCAGCTACAAAGTTTGTCTTCATTCATCACCCTTCTCTCAAGCAAGTGA 2040  
Db 1981 GAGTTTTCAGCTACAAAGTTTGTCTTCATTCATCACCCTTCTCTCAAGCAAGTGA 2040  
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAG 2100  
Db 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAG 2100  
QY 2101 TACACATCCAAATTTATGCAAAATCTGCAACGTTGATTTTACTGTGACAAACATGACTT 2160  
Db 2101 TACACATCCAAATTTATGCAAAATCTGCAACGTTGATTTTACTGTGACAAACATGACTT 2160  
QY 2161 TATACTGAGCGCTCGCCCAATTTGGCACCCGTTACTTACCGTCCCTGT 2209  
Db 2161 TATACTGAGCGCTCGCCCAATTTGGCACCCGTTACTTACCGTCCCTGT 2209  
RESULT 2  
AD227052  
ID AD227052 standard; DNA; 2211 BP.  
XX  
AC AD227052;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE Adeno-associated virus DNA SEQ ID NO 202.  
XX  
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;  
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;  
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;  
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;  
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;  
KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.  
XX  
OS Adeno-associated virus.  
XX  
PN W02005033321-A2.  
XX  
PD 14-APR-2005.  
XX  
PF 30-SEP-2004; 2004MO-US028817.  
XX  
PR 30-SEP-2003; 2003US-0508226P.  
XX  
PR 29-APR-2004; 2004US-0566546P.  
XX  
PR (UYPE-) UNIV PENNSYLVANIA.  
XX  
PI Wilson JM, Gao G, Alvira MR, Vandenbergh LH;  
PI WPI; 2005-285437/29.  
DR  
DR New adeno-associated virus (AAV) clade comprising at least three AAV  
PT members, useful for preventing and/or treating arthritis, multiple  
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial  
PT infection and cancer.  
PT  
XX Disclosure; SEQ ID NO 202; 569pp; English.  
PS  
XX The invention relates to an adeno-associated virus (AAV) clade comprising  
XX at least three AAV members, where each member of the AAV clade is  
XX phylogenetically related as determined using a Neighbor-Joining heuristic  
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson  
XX correction distance measurement of no more than 0.05. The methods and  
XX compositions of the present invention are useful for the prevention  
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,  
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's  
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative

CC colitis. The present sequence represents an adeno-associated virus DNA.  
XX  
SQ Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2209; DB 14; Length 2211;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTGCGGATGTTATCTTCCAGATTGGCTCGAGGCAACCTCTCTGAGGCAATTCGC 60  
Db 1 ATGGCTGCGGATGTTATCTTCCAGATTGGCTCGAGGCAACCTCTCTGAGGCAATTCGC 60  
QY 61 GAGTGTGGGACTTGAACCTGGAGCCCGAAGCCCAAGCCCAACGAGCAAAAGCGAGAC 120  
Db 61 GAGTGTGGGACTTGAACCTGGAGCCCGAAGCCCAAGCCCAACGAGCAAAAGCGAGAC 120  
QY 121 GACGCCGGGGTCTGGTCTCTTCCAGTACAGTACCTCGGACCCCTTCAACGAGCTCGAC 180  
Db 121 GACGCCGGGGTCTGGTCTCTTCCAGTACAGTACCTCGGACCCCTTCAACGAGCTCGAC 180  
QY 181 AAGGGGAGCCCGTCAACGCGCGGACGCGAGCGGCGCTCGAGCAGCAAGGCTTACGAC 240  
Db 181 AAGGGGAGCCCGTCAACGCGCGGACGCGAGCGGCGCTCGAGCAGCAAGGCTTACGAC 240  
QY 241 CAGCAGCTCAAGCGGGTGACAAATCGGTACCTGCGGTATAACCAACGCGCGGAGTTT 300  
Db 241 CAGCAGCTCAAGCGGGTGACAAATCGGTACCTGCGGTATAACCAACGCGCGGAGTTT 300  
QY 301 CAGGAGCGTCTCAAGAGATACGCTCTTTTGGGGCAACCTCGGGGAGCAGTCTTCCAG 360  
Db 301 CAGGAGCGTCTCAAGAGATACGCTCTTTTGGGGCAACCTCGGGGAGCAGTCTTCCAG 360  
QY 361 GCCAAGAAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGCGGCTCT 420  
Db 361 GCCAAGAAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGCGGCTCT 420  
QY 421 GGAAGAAAGCTCCGGTAGAGCAGTGCACCAAGAGCCAGACTCTCTCTCGGCAATCGGC 480  
Db 421 GGAAGAAAGCTCCGGTAGAGCAGTGCACCAAGAGCCAGACTCTCTCTCGGCAATCGGC 480  
QY 481 AAGACAGGCGCAGCGCCGCTAAAGAGACTCAATTTTGGTTCAGACTGCGGACTCAGAG 540  
Db 481 AAGACAGGCGCAGCGCCGCTAAAGAGACTCAATTTTGGTTCAGACTGCGGACTCAGAG 540  
QY 541 TCAGTCCCCGATCCCAACCTCTCGGAGAACCTTCCAGCAACCCCGCTCTGTGGGACCT 600  
Db 541 TCAGTCCCCGATCCCAACCTCTCGGAGAACCTTCCAGCAACCCCGCTCTGTGGGACCT 600  
QY 601 ACTACAAATGGCTTCAGGCGGTGGCGCAACATGGGAGCAATTAACGAGCGCGGACGGA 660  
Db 601 ACTACAAATGGCTTCAGGCGGTGGCGCAACATGGGAGCAATTAACGAGCGCGGACGGA 660  
QY 661 GTGGGTAAATGCTTCAGGAAATTTGGCAATTCGATTTCCATGCTGGGCGGACAGTATC 720  
Db 661 GTGGGTAAATGCTTCAGGAAATTTGGCAATTCGATTTCCATGCTGGGCGGACAGTATC 720  
QY 721 ACCACAGCAACCGCACCTTGGCCACCTTCCAGCAACCTTCAATAAACCACTCTTACAGCAATC 780  
Db 721 ACCACAGCAACCGCACCTTGGCCACCTTCCAGCAACCTTCAATAAACCACTCTTACAGCAATC 780  
QY 781 TCCAGTGTCTTCAACGGGGGCTCGGCGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 840  
Db 781 TCCAGTGTCTTCAACGGGGGCTCGGCGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 840  
QY 841 GGGTATTTGATTTCAACAGATTCACCTGCGCTTTTCCAGCTGCTGGGAGCGGACTC 900  
Db 841 GGGTATTTGATTTCAACAGATTCACCTGCGCTTTTCCAGCTGCTGGGAGCGGACTC 900  
QY 901 ATCAACAACTTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960  
Db 901 ATCAACAACTTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960  
QY 961 GTCAAGGAGGTTCACGAGCAATGATGCGGTCAACCATCGCTTAATTAACCTTACAGCAGC 1020

Db 961 GTACAGGAGGTACGAGCAATGATGGCGTCACAAACCATCGCTAATAACCTTACCAGCAG 1020  
Qy 1021 GTTCAAGTCTTCGAGCTCGAGATACAGAGTTCGATCGTCTCGCTCGCTCGGACACAG 1080  
Db 1021 GTTCAAGTCTTCGAGCTCGAGATACAGAGTTCGATCGTCTCGCTCGCTCGGACACAG 1080  
Qy 1081 GGTGCTCCCTCCGTTCCCGGGGAGCGTTCATGATTCGCAATACGGCTACCTGAGC 1140  
Db 1081 GGTGCTCCCTCCGTTCCCGGGGAGCGTTCATGATTCGCAATACGGCTACCTGAGC 1140  
Qy 1141 CTCAACAATGGCAGCCGAGCGTGGAGCGTTTCATCTCTTTTACTGCTGGAATATTTCCCT 1200  
Db 1141 CTCAACAATGGCAGCCGAGCGTGGAGCGTTTCATCTCTTTTACTGCTGGAATATTTCCCT 1200  
Qy 1201 TCTCAGATCGTGAGAAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 1260  
Db 1201 TCTCAGATCGTGAGAAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT 1260  
Qy 1261 TTCCACAGCAGCTACGCGCACAGCCAGAGCGCTGGACCGGCTGATGATCTCTCATCGAC 1320  
Db 1261 TTCCACAGCAGCTACGCGCACAGCCAGAGCGCTGGACCGGCTGATGATCTCTCATCGAC 1320  
Qy 1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAAGGAC 1380  
Db 1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAAGGAC 1380  
Qy 1381 TTGCTGTTTAGCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAAACCTGGCTACCT 1440  
Db 1381 TTGCTGTTTAGCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAAACCTGGCTACCT 1440  
Qy 1441 GGACCTGTATTCCGAGCAGCGCGTTCCTTAAACAAACAAACAGACAAACAGCAAT 1500  
Db 1441 GGACCTGTATTCCGAGCAGCGCGTTCCTTAAACAAACAAACAGACAAACAGCAAT 1500  
Qy 1501 TTACCTGGAGCTGGTCTTCAAAATATAACCTCAATGGCGTGAATCCATCAACCT 1560  
Db 1501 TTACCTGGAGCTGGTCTTCAAAATATAACCTCAATGGCGTGAATCCATCAACCT 1560  
Qy 1561 GGCACCTGTATGCGCTCACAAAGACGACGAAGCAAGTCTTTCCCATGAGCGGTGTC 1620  
Db 1561 GGCACCTGTATGGCTCACAAAGACGACGAAGCAAGTCTTTCCCATGAGCGGTGTC 1620  
Qy 1621 ATGATTTTGGAAAAGAGCGCGGAGCTTCAACACATGCATTTGGACATGTCATGATT 1680  
Db 1621 ATGATTTTGGAAAAGAGCGCGGAGCTTCAACACATGCATTTGGACATGTCATGATT 1680  
Qy 1681 ACAGACGAAGAGAAATTAAGCCACTAACCTCTGGCCACCGAAAGATTTGGACCGTG 1740  
Db 1681 ACAGACGAAGAGAAATTAAGCCACTAACCTCTGGCCACCGAAAGATTTGGACCGTG 1740  
Qy 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTCATGCTATGGGA 1800  
Db 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTCATGCTATGGGA 1800  
Qy 1801 GCATTAACCTGGCATGTGTGGCAAGATAGAGCTGTACTCGAGAGTCCCATTTGGGCC 1860  
Db 1801 GCATTAACCTGGCATGTGTGGCAAGATAGAGCTGTACTCGAGAGTCCCATTTGGGCC 1860  
Qy 1861 AAAATTCCTCAACAGATGGACACTTTTCAACCCGTCTCTTATGGCGCGCTTTGACATC 1920  
Db 1861 AAAATTCCTCAACAGATGGACACTTTTCAACCCGTCTCTTATGGCGCGCTTTGACATC 1920  
Qy 1921 AGAAACCGCGCTCTCAGATCTCATCAAAACACAGCGCTGTTCCTCGAATCTCCGGCG 1980  
Db 1921 AGAAACCGCGCTCTCAGATCTCATCAAAACACAGCGCTGTTCCTCGAATCTCCGGCG 1980  
Qy 1981 GAGTTTTTCAAGTTCAGATTTGCTTCATTCATCAACCAATCTCCACAGCAAGTGAAT 2040  
Db 1981 GAGTTTTTCAAGTTCAGATTTGCTTCATTCATCAACCAATCTCCACAGCAAGTGAAT 2040  
Qy 2041 GTGGAATTTGAATGGGAGCTGCAGAAAGAAAAACGAAGCGCTGGGAATCCCCGAAGTGCAG 2100

Db 2041 GTGGAATTTGAATGGGAGCTGCAGAAAGAAAAACGAAGCGCTGGGAATCCCGAAGTGCAG 2100  
Qy 2101 TACACATCCAATTTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACATGACTT 2160  
Db 2101 TACACATCCAATTTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACATGACTT 2160  
Qy 2161 TATACGTAGCCTCGGCCCATTTGGCAACCGTTTACCTTACCCGTCCTCTGT 2209  
Db 2161 TATACGTAGCCTCGGCCCATTTGGCAACCGTTTACCTTACCCGTCCTCTGT 2209

RESULT 3  
AD226929  
ID AD226929 standard; DNA; 2211 BP.  
XX  
AC AD226929;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE Adeno-associated virus DNA SEQ ID NO 79.  
XX  
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;  
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;  
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;  
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;  
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;  
KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.  
XX  
OS Adeno-associated virus.  
XX  
PN WO2005033321-A2.  
XX  
PD 14-APR-2005.  
XX  
PF 30-SEP-2004; 2004WO-US028817.  
XX  
PR 30-SEP-2003; 2003US-0508226P.  
PR 29-APR-2004; 2004US-0566546P.  
XX  
PA (UYPE-) UNIV PENNSYLVANIA.  
XX  
PI Wilson JM, Gao G, Alvira MR, Vandenbergh LH;  
XX  
WPI; 2005-285437/29.  
XX  
PT New adeno-associated virus (AAV) clade comprising at least three AAV  
members, useful for preventing and/or treating arthritis, multiple  
sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial  
infection and cancer.  
XX  
PS Disclosure; SEQ ID NO 79; 569pp; English.  
XX  
CC The invention relates to an adeno-associated virus (AAV) clade comprising  
at least three AAV members, where each member of the AAV clade is  
phylogenetically related as determined using a Neighbor-Joining heuristic  
by a bootstrap value of at least 75 % per 1000 isolates and a Poisson  
correction distance measurement of no more than 0.05. The methods and  
compositions of the present invention are useful for the prevention  
and/or treatment of rheumatoid arthritis, multiple sclerosis,  
sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's  
disease, hemophilia, HIV, bacterial infection, cancer and ulcerative  
colitis. The present sequence represents an adeno-associated virus DNA.  
XX  
SQ Sequence 2211 BP; 577 A; 652 C; 538 G; 444 T; 0 U; 0 Other;

Query Match 100.0%; Score 2209; DB 14; Length 2211;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60  
Db 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60

Qy	61	GAGTGTGGGACTTTGAAACCTCTGGAGCCCCGAAAGCCCAAGCCAAACGAGAAAAGCAGGAC	120
Db	61	GAGTGTGGGACTTTGAAACCTCTGGAGCCCCGAAAGCCCAAGCCAAACGAGAAAAGCAGGAC	120
Qy	121	GAGGGCCGGGCTCTGGTGTCTCTCTGGCTCAAGTACTCTCGAGCCCTTCAACGACTCGAC	180
Db	121	GAGGGCCGGGCTCTGGTGTCTCTCTGGCTCAAGTACTCTCGAGCCCTTCAACGACTCGAC	180
Qy	181	AAGGGGAGCCCGTCAACGCGGGGAGCGCAGCGGCCCTCGAGCACGAAGGCTTACGAC	240
Db	181	AAGGGGAGCCCGTCAACGCGGGGAGCGCAGCGGCCCTCGAGCACGAAGGCTTACGAC	240
Qy	241	CAGCAGCTCAAGCGGGGTGAACAATCCGTACTCTGGGTATTAACCAAGCCGACCGAGTTT	300
Db	241	CAGCAGCTCAAGCGGGGTGAACAATCCGTACTCTGGGTATTAACCAAGCCGACCGAGTTT	300
Qy	301	CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
Db	301	CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
Qy	361	GCCAAAGAGCGGGTCTTCGAAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGGTCTCT	420
Db	361	GCCAAAGAGCGGGTCTTCGAAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGGTCTCT	420
Qy	421	GGAAAGAAACGTCGCGTAGACAGTCGCCCAAGAGCCAGACTCTCTCGGGCATCGGC	480
Db	421	GGAAAGAAACGTCGCGTAGACAGTCGCCCAAGAGCCAGACTCTCTCTCGGGCATCGGC	480
Qy	481	AAGCAGGCCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTTCAGACTGGGACTCGAG	540
Db	481	AAGCAGGCCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTTCAGACTGGGACTCGAG	540
Qy	541	TCAGTCCCGCATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGTGTGGGACCT	600
Db	541	TCAGTCCCGCATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGTGTGGGACCT	600
Qy	601	ACTACAATGGCTTCAGCGGTGGCGCACCAATATGCGACACAATAACGAAGGGCGCGACGGA	660
Db	601	ACTACAATGGCTTCAGCGGTGGCGCACCAATATGCGACACAATAACGAAGGGCGCGACGGA	660
Qy	661	GTGGTAAATGGCTCAGGAATTTGGCATTTGGATTTCCACATGGCTGGGGCAGCAGTCAATC	720
Db	661	GTGGTAAATGGCTCAGGAATTTGGCATTTGGATTTCCACATGGCTGGGGCAGCAGTCAATC	720
Qy	721	ACCACAGCACCCGACCTGGGCCCTTGCCACCTTACAAATTAACACCTCTTACAAGCAATC	780
Db	721	ACCACAGCACCCGACCTGGGCCCTTGCCACCTTACAAATTAACACCTCTTACAAGCAATC	780
Qy	781	TCCAGTGTCTTCAACGGGGGCGAGCAACGACCAACACTACTTCGGCTACAGCACCCCTGG	840
Db	781	TCCAGTGTCTTCAACGGGGGCGAGCAACGACCAACACTACTTCGGCTACAGCACCCCTGG	840
Qy	841	GGGTATTTTGAATTTCAACAGATTCACATGCCACTTTTTCACCACTGACTGGCAGCGACTC	900
Db	841	GGGTATTTTGAATTTCAACAGATTCACATGCCACTTTTTCACCACTGACTGGCAGCGACTC	900
Qy	901	ATCAACAACAATTCGGGATTCGGGCCCAAGAGACTCAACTTTCAAATCTTTCAACATCAA	960
Db	901	ATCAACAACAATTCGGGATTCGGGCCCAAGAGACTCAACTTTCAAATCTTTCAACATCAA	960
Qy	961	GTCAAGGAGGTTCAGCAGCAATGATGGGCTCAACAACCTATGCTAATAACCTTTACGACACG	1020
Db	961	GTCAAGGAGGTTCAGCAGCAATGATGGGCTCAACAACCTATGCTAATAACCTTTACGACACG	1020
Qy	1021	GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCGGTACGTCTCGGCTCTGGCAGCCAG	1080
Db	1021	GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCGGTACGTCTCGGCTCTGGCAGCCAG	1080
Qy	1081	GGCTGCCCTCCCTCGTTCGCCGCGGACGTGTTTCAATGTTCCGCAATACGGCTACTCTGACG	1140
Db	1081	GGCTGCCCTCCCTCGTTCGCCGCGGACGTGTTTCAATGTTCCGCAATACGGCTACTCTGACG	1140
Qy	1141	CTCAACAATGCGCAGCAAGCGCTGGGACGTTTCATCTCTTTTACTGCTGGAATATTTCCCT	1200

[illegible]





Db 3461 GGACCCCTGTTATTCGGCAGCAGCGCGTTCTTAAACAAAAACAGACAAACACAGCAAT 3520  
QY 1501 TTTTACCTGGACGTGGTCTTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCT 1560  
Db 3521 TTTTACCTGGACGTGGTCTTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCT 3580  
QY 1561 GGCATCTGCTATGGCCTTCACACAAAGACGACGAAGCAAGTTCTTTCCCATGAGCGGTGTC 1620  
Db 3581 GGCATCTGCTATGGCCTTCACACAAAGACGACGAAGCAAGTTCTTTCCCATGAGCGGTGTC 3640  
QY 1621 ATGATTTTGGAAAGAGAGCGCGGAGCTTCACACACATGCAATGGACAAATGTCATGATT 1680  
Db 3641 ATGATTTTGGAAAGAGAGCGCGGAGCTTCACACACATGCAATGGACAAATGTCATGATT 3700  
QY 1681 ACAGACGAAGAGAAATTAAGCCACTAAACCTGTGGCCACCGCAAGAGATTTGGGACCGTG 1740  
Db 3701 ACAGACGAAGAGAAATTAAGCCACTAAACCTGTGGCCACCGCAAGAGATTTGGGACCGTG 3760  
QY 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACACACCTTCGACCGGAGATGTGCATGTATGGGA 1800  
Db 3761 GCAGTCAATTTCCAGAGCAGCAGCAGACACACCTTCGACCGGAGATGTGCATGTATGGGA 3820  
QY 1801 GCATTAACCTGGATGCTGTGGCAAGATAGAGAGCTGACTGCAGGGTCCCATTTGGGCC 1860  
Db 3821 GCATTAACCTGGATGCTGTGGCAAGATAGAGAGCTGACTGCAGGGTCCCATTTGGGCC 3880  
QY 1861 AAAATTCCTCACACAGATGGACACTTTACCCGCTCTCTTATGGCGGCTTTGGACTC 1920  
Db 3881 AAAATTCCTCACACAGATGGACACTTTACCCGCTCTCTTATGGCGGCTTTGGACTC 3940  
QY 1921 AAGAACCCGCTCCTCAGATCCTCATCAAAAAACACGCCCTGTTCTTCGCAATCTCCCGCG 1980  
Db 3941 AAGAACCCGCTCCTCAGATCCTCATCAAAAAACACGCCCTGTTCTTCGCAATCTCCCGCG 4000  
QY 1981 GAGTTTTCAGTCAAAAGTTTGCTTCAATCATCACCACCAATCTCCACAGACAAAGTGA 2040  
Db 4001 GAGTTTTCAGTCAAAAGTTTGCTTCAATCATCACCACCAATCTCCACAGACAAAGTGA 4060  
QY 2041 GTGGAATTAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAG 2100  
Db 4061 GTGGAATTAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAG 4120  
QY 2101 TACACATCAATTAATGCAAAATTCGCCAAGCTTGATTTTACTGTGGACAAACATGACTT 2160  
Db 4121 TACACATCAATTAATGCAAAATTCGCCAAGCTTGATTTTACTGTGGACAAACATGACTT 4180  
QY 2161 TATACTGAGCTCGCCCAATGGCACCCGTTACCTTACCGTCCCTGT 2209  
Db 4181 TATACTGAGCTCGCCCAATGGCACCCGTTACCTTACCGTCCCTGT 4229

RESULT 5  
ID AD246594  
XX AD246594 standard; DNA; 4347 BP.  
AC AD246594;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE HSV-AAV sequence rep2cap1 for AAV/HSV vectors.  
XX  
KW Genetic engineering; gene therapy; vector; cancer; neoplasm; cytostatic;  
KW respiratory disease; respiratory-gen.; neurological disease;  
KW neuroprotective; db.  
XX  
OS Adeno-associated virus.  
OS Human herpesvirus 1.  
XX  
PN WO2005035743-A1.  
XX  
PD 21-APR-2005.  
XX  
PF 15-OCT-2003; 2003WO-CN000861.

XX 15-OCT-2003; 2003WO-CN000861.  
PR (AGTC-) AGTC GENE TECHNOLOGY CO LTD.  
XX  
XX Wu X, Cao H, Dong X;  
PI WPI; 2005-296279/30.  
XX  
DR Large-scale production, isolation and purification of serotype adeno-  
XX associated virus vectors by infecting cells with recombinant herpes  
PT simplex virus vectors, for use in gene therapy of e.g. cancer and  
PT respiratory diseases.  
XX  
PS Claim 1; SEQ ID NO 1; 79pp; Chinese.  
XX  
CC The invention relates to a recombinant herpes simplex virus (HSV)  
CC characterized in that its genome is inserted with a DNA sequence selected  
CC from nucleotide sequences of AD246594-AD246598 or their homologous  
CC sequences. Also included are preparing the recombinant HSV (comprising  
CC the construction of DNA fragments that contains sequences AD246594-  
CC AD246598, and respectively inserting these 5 DNA fragments into genome of  
CC HSV by applying genetic engineering to give the recombinant HSV), large-  
CC scale production of 5 serotype recombinant adeno-associated viruses (AAV)  
CC 1, 3, 4, 5, and 6 (by preparing the recombinant HSV, establishing a one-  
CC strain vector cell i.e. recombinant AAV vector cell-line, infecting the  
CC recombinant AAV under the infective effect of the 5 recombinant HSV in  
CC the vector cell-line, separating and purifying the 5 serotype recombinant  
CC AAV after lysing the AAV-containing cells and culture liquor to give a  
CC crude lysate and further purification of rAAV by density-gradient  
CC centrifugation or affinity chromatography) and a recombinant vector  
CC plasmid pSNV-NX (containing ITR at both ends of AAV-1, AAV-3, AAV-4, AAV  
CC -5 or AAV-6 genome, with immediate-early enhancer and promoter of  
CC cytomegalovirus, and a polyclonal site and a polyA signal, respectively,  
CC between the ITR (inverted terminal repeat), and neomycin-resistance gene-  
CC expressing cassette at outer edge of ITR). The DNA sequence is inserted  
CC into XbaI site in UL2 or UL44 gene in HSV genome. The DNA sequence of  
CC sequence AD246594-AD246598 can also be inserted into other non-essential  
CC gene regions in HSV genome. The recombinant HSV is optionally inserted  
CC with other DNA sequences homologous with the already-specified fragments.  
CC The vectors are for use in gene therapy of e.g. cancer, respiratory  
CC diseases and neural diseases. The virus vectors are safe, with long  
CC expression time and wide-spectrum of cell infection, even non-cleaved  
CC cells and reverse axonal conduction through the incorporated HSV vectors,  
CC and high transfer efficiently. The present sequence is the HSV-AAV  
CC sequence rep2cap1 for the AAV/HSV vectors of the invention.  
XX  
SQ Sequence 4347 BP; 1106 A; 1186 C; 1142 G; 913 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2209; DB 14; Length 4347;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCGATTGCG 60  
Db 2021 ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCGATTGCG 2080  
QY 61 GAGTGTGGGACTTGAACCTCGAGCCCGAGCCCAAGCCAAACAGCAAGAGCAGGAC 120  
Db 2081 GAGTGTGGGACTTGAACCTCGAGCCCGAGCCCAAGCCAAACAGCAAGAGCAGGAC 2140  
QY 121 GACGGCCGGGCTCTGGTGTCTTCTGGCTACAAAGTACTCGGACCTTCAACGACTCGAC 180  
Db 2141 GACGGCCGGGCTCTGGTGTCTTCTGGCTACAAAGTACTCGGACCTTCAACGACTCGAC 2200  
QY 181 AAGGGGGAGCCCGTCAACGCGCGGACGCGAGCGGCCCTCGAGCACCAAGGCTTACGAC 240  
Db 2201 AAGGGGGAGCCCGTCAACGCGCGGACGCGAGCGGCCCTCGAGCACCAAGGCTTACGAC 2260  
QY 241 CAGCAGCTCAAGCGGGTGACAAATCCGTACCTGCGGTATTAACACCCGCGGAGTTT 300  
Db 2261 CAGCAGCTCAAGCGGGTGACAAATCCGTACCTGCGGTATTAACACCCGCGGAGTTT 2320



QY 301 CAGGAGCGTCTGCAAGATAGTCTTTTGGGGCAACCTCGGGGAGCAGCTCTCCAG 360  
Db 2321 CAGGAGCGTCTGCAAGATAGTCTTTTGGGGCAACCTCGGGGAGCAGCTCTCCAG 2380  
QY 361 GCCAAGACGGGTCTCGAACCCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT 420  
Db 2381 GCCAAGACGGGTCTCGAACCCTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT 2440  
QY 421 GGAAAGAAACGTCCGGTAGAGAGTCCGCCAAGAGCCAGACTCTCTCTCGGGCATCGGC 480  
Db 2441 GGAAAGAAACGTCCGGTAGAGAGTCCGCCAAGAGCCAGACTCTCTCTCGGGCATCGGC 2500  
QY 481 AAGACAGGCGAGCAGCCCTTAAAGAGACTCAATTTTGGTCAGAGTGGCGACTCAGAG 540  
Db 2501 AAGACAGGCGAGCAGCCCTTAAAGAGACTCAATTTTGGTCAGAGTGGCGACTCAGAG 2560  
QY 541 TCAGTCCCCGATCCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGGGACCT 600  
Db 2561 TCAGTCCCCGATCCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGGGACCT 2620  
QY 601 ACTAATAGCTTTCAGCGGTGGCGCAACCAATGGCAGACAAATAACGAAGCGCGCAGCGA 660  
Db 2621 ACTAATAGCTTTCAGCGGTGGCGCAACCAATGGCAGACAAATAACGAAGCGCGCAGCGA 2680  
QY 661 GTGGGTAAAGCTTCAGGAATTTGGCAATTTGGCAATTCACATGGCTGGGCGACAGAGTCATC 720  
Db 2681 GTGGGTAAAGCTTCAGGAATTTGGCAATTTGGCAATTCACATGGCTGGGCGACAGAGTCATC 2740  
QY 721 ACCACAGACCCCGACCTTGGGCTTTGCCACCTTCAATAACCACTCTTCAAGCAAAATC 780  
Db 2741 ACCACAGACCCCGACCTTGGGCTTTGCCACCTTCAATAACCACTCTTCAAGCAAAATC 2800  
QY 781 TCAGTGTCTTAACGGGGCGCAGCAACCACTAATTTCGGCTACAGCAACCCCGTGG 840  
Db 2801 TCAGTGTCTTAACGGGGCGCAGCAACCACTAATTTCGGCTACAGCAACCCCGTGG 2860  
QY 841 GGGTATTTTGAATTTCAACAGATTCACCTGCCACTTTTACCAGCTGAGTGGCAGGACTC 900  
Db 2861 GGGTATTTTGAATTTCAACAGATTTCCACTGCCACTTTTACCAGCTGAGTGGCAGGACTC 2920  
QY 901 ATCAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA 960  
Db 2921 ATCAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA 2980  
QY 961 GTCAAGGAGTCAAGCAATGATGGCGTCAAAACATCGCTTAATTAACCTTACAGCAG 1020  
Db 2981 GTCAAGGAGTCAAGCAATGATGGCGTCAAAACATCGCTTAATTAACCTTACAGCAG 3040  
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAAGTCTCGGCTCTGGCACAG 1080  
Db 3041 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCGTAAGTCTCGGCTCTGGCACAG 3100  
QY 1081 GGCTGCTCTCTCGGCTTCCGGCGGAGCTGTTTCATGATTTCCGCAATACGGCTACCTGAGC 1140  
Db 3101 GGCTGCTCTCTCGGCTTCCGGCGGAGCTGTTTCATGATTTCCGCAATACGGCTACCTGAGC 3160  
QY 1141 CTCAACAAATGGCAGCAAGCGTGGACGTTTCATCTCTTACTGCTGGGAATATTTCCTCT 1200  
Db 3161 CTCAACAAATGGCAGCAAGCGTGGACGTTTCATCTCTTACTGCTGGGAATATTTCCTCT 3220  
QY 1201 TCTCAGATGCTGAGAACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAGTGCCT 1260  
Db 3221 TCTCAGATGCTGAGAACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAGTGCCT 3280  
QY 1261 TTCCACAGAGCTTACCGGCACAGCCTGGAGCGGTGATGATCTCTCTCATCGAC 1320  
Db 3281 TTCCACAGAGCTTACCGGCACAGCCTGGAGCGGTGATGATCTCTCTCATCGAC 3340  
QY 1321 CAATACCTGTATTACTTGAAACAGAACTCAAAATCACTCGGAAGTCCCAAAACAAAGGAC 1380  
Db 3341 CAATACCTGTATTACTTGAAACAGAACTCAAAATCACTCGGAAGTCCCAAAACAAAGGAC 3400

QY 1381 TTGCTGTTTAGCGGTGCTCCAGCTGGCATGTCTGTTCCAGCCCAAAAACTGGCTACT 1440  
Db 3401 TTGCTGTTTAGCGGTGCTCCAGCTGGCATGTCTGTTCCAGCCCAAAAACTGGCTACT 3460  
QY 1441 GGACCTCTTATCGGCAGCAGCGCTTTCTAAAAAAGAGCAAGCAAAACAAACGCAAT 1500  
Db 3461 GGACCTCTTATCGGCAGCAGCGCTTTCTAAAAAAGAGCAAGCAAAACAAACGCAAT 3520  
QY 1501 TTTACTCGGACTGTGGTCTTCAAAATATAAAGCTCAATGGGGTGAATCATCATCAACCT 1560  
Db 3521 TTTACTCGGACTGTGGTCTTCAAAATATAAAGCTCAATGGGGTGAATCATCATCAACCT 3580  
QY 1561 GGCACGTATATGGCTCTCACAAAGACGACGAAGCAAGTCTTTTCCCATGAGCGGTGC 1620  
Db 3581 GGCACGTATATGGCTCTCACAAAGACGACGAAGCAAGTCTTTTCCCATGAGCGGTGC 3640  
QY 1621 ATGATTTTGGAAAAAGAGCGCGGAGCTTCAAAACACTGCAATTTGGCAAAATGCTAT 1680  
Db 3641 ATGATTTTGGAAAAAGAGCGCGGAGCTTCAAAACACTGCAATTTGGCAAAATGCTAT 3700  
QY 1681 ACAGACGAAGAGGAAATTAAGGCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740  
Db 3701 ACAGACGAAGAGGAAATTAAGGCACTAAACCTGTGGCCACCGAAAGATTTGGGACCGTG 3760  
QY 1741 GCAGTCAATTTCCAGAGCAGCAGACAGACCTCGACCGGAGATGTGCATGTATGGGA 1800  
Db 3761 GCAGTCAATTTCCAGAGCAGCAGACAGACCTCGACCGGAGATGTGCATGTATGGGA 3820  
QY 1801 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCAATTTGGGCC 1860  
Db 3821 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCAATTTGGGCC 3880  
QY 1861 AAAATTCCTCACAGATGGACACTTTTCAACCGCTCTCTTTATGGGCGGCTTTGGACTC 1920  
Db 3881 AAAATTCCTCACAGATGGACACTTTTCAACCGCTCTCTTTATGGGCGGCTTTGGACTC 3940  
QY 1921 AAGAACCCGCTCTCAGATCTCATCAAAACACGCTGTTTCCGCAATCCTCCGGCG 1980  
Db 3941 AAGAACCCGCTCTCAGATCTCATCAAAACACGCTGTTTCCGCAATCCTCCGGCG 4000  
QY 1981 GAGTTTTCAGCTACAAAGTTTGTCTTATTCATCAACCAATCTCCACAGGCAAGTGAGT 2040  
Db 4001 GAGTTTTCAGCTACAAAGTTTGTCTTATTCATCAACCAATCTCCACAGGCAAGTGAGT 4060  
QY 2041 GTGGAAAATTGAATGGGAGCTGCAGAAAAGAAAGACGCTGGAATCCCGAAGTGAC 2100  
Db 4061 GTGGAAAATTGAATGGGAGCTGCAGAAAAGAAAGACGCTGGAATCCCGAAGTGAC 4120  
QY 2101 TACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGACATT 2160  
Db 4121 TACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGACATT 4180  
QY 2161 TATACTGAGCCTCGCCCATTTGGCACCGGTTACCTTTACCGTCCCTCTGT 2209  
Db 4181 TATACTGAGCCTCGCCCATTTGGCACCGGTTACCTTTACCGTCCCTCTGT 4229

RESULT 6  
AAD00772  
ID AAD00772 standard; DNA; 4718 BP.  
XX  
AC AAD00772;  
XX  
DT 08-SEP-2000 (first entry)  
XX  
DE Adeno-associated virus serotype 1 DNA.  
XX  
KW Adeno-associated virus serotype 1; AAV-1; rep protein; capsid protein;  
XX cap protein; recombinant viral vector; gene delivery; gene therapy;  
XX vaccine; transgene; ss.  
OS Adeno associated virus serotype 1.

```

FH      Location/Qualifiers
FT      1. .143
FT      /*tag= a
FT      /label= 5' ITR
FT      /note= "Inverted terminal repeat which is capable of
FT      forming T-shaped hairpin structure"
FT      89. .110
FT      /*tag= b
FT      /bound moiety= "Rep protein"
FT      124. .125
FT      /tag= c
FT      /note= "terminal resolute site (TRS)"
FT      219. .226
FT      /*tag= d
FT      /bound moiety= "USF"
FT      /note= "g box"
FT      236. .299
FT      /*tag= e
FT      /label= P5_promoter
FT      237. .245
FT      /*tag= f
FT      /bound moiety= "YY1 factor"
FT      270. .275
FT      /tag= g
FT      /label= P5_TATA-Box
FT      299. .306
FT      /*tag= h
FT      /note= "YY1/p5 RNA"
FT      335. .2272
FT      /*tag= j
FT      /product= "Rep 68"
FT      /function= "regulates replication and integration of AAV
FT      DNA into host cell's chromosome"
FT      /note= "the coding region is interrupted by intron"
FT      335. .2206
FT      /*tag= i
FT      /product= "Rep 78"
FT      /function= "regulates replication and integration of AAV
FT      DNA into host cell's chromosome"
FT      857. .862
FT      /*tag= l
FT      /label= P19_TATA_Box
FT      882. .883
FT      /*tag= m
FT      /note= "P19 RNA"
FT      1007. .2272
FT      /*tag= o
FT      /product= "Rep 40"
FT      /function= "regulates replication and integration of AAV
FT      DNA into host cell's chromosome"
FT      /note= "the coding region is interrupted by intron"
FT      1007. .2206
FT      /*tag= n
FT      /product= "Rep 52"
FT      /function= "regulates replication and integration of AAV
FT      DNA into host cell's chromosome"
FT      1836. .1841
FT      /*tag= p
FT      /label= P40_TATA-BOX
FT      1875. .1876
FT      /*tag= q
FT      /note= "P40 RNA"
FT      1924. .2220
FT      /*tag= k
FT      /note= "This region interrupts the coding sequence of Rep
FT      68 and Rep 40"
FT      2223. .4433
FT      /*tag= r
FT      /product= "vpi protein"
FT      /note= "Capsid protein"
FT      2634. .4433
FT      /*tag= s
FT      /product= "VP2 protein"

```

```

FT      /note= "Capsid protein"
FT      /partial
FT      2829. .4433
FT      /*tag= t
FT      /product= "vp3 protein"
FT      /note= "Capsid protein"
FT      4447. .4452
FT      /*tag= u
FT      repeat_unit
FT      4576. .4718
FT      /*tag= v
FT      /label= 3' ITR
FT      /note= "Inverted terminal repeat which is capable of
FT      forming T-shaped hairpin structure"
FT      WO200028061-A2.
FT      18-MAY-2000.
FT      02-NOV-1999; 99WO-US025694.
FT      05-NOV-1998; 98US-0107114P.
FT      (UYPE-) UNIV PENNSYLVANIA.
FT      Wilson JM, Xiao W;
FT      WPI; 2000-376571/32.
FT      P-PSDB; AAY71161, AAY71164, AAY71165, AAY71166, AAY71167, AAY71168,
FT      AAY71169.
FT      Novel adeno-associated virus serotype 1 polynucleotide useful for
FT      preparation of medicament for delivery of a transgene to a host.
FT      Claim 1; Fig 1; 108pp; English.
FT      The present sequence is an adeno-associated virus serotype 1 (AAV-1) DNA
FT      characterised by two inverted terminal repeats (ITR) and open reading
FT      frames for rep and capsid (cap) proteins. The rep reading frame encodes
FT      four proteins, Rep 78, Rep 68, Rep 52 and Rep 40, while the cap encoding
FT      frame encodes three structural proteins, VP1, VP2 and VP3. The AAV-1
FT      sequence or its fragments particularly ITRs, rep and cap coding regions,
FT      are useful in production of recombinant viral vectors for gene delivery.
FT      These vectors can be used as gene therapy vectors, vaccine vectors or
FT      antisense delivery vectors. The AAV-1 does not induce the formation of
FT      neutralising antibodies specific to any serotype of AAV hence is useful
FT      for transforming host cells, and in the preparation of a medicament for
FT      the delivery of transgene to a host
FT      Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;
FT      Query Match 100.0%; Score 2209; DB 3; Length 4718;
FT      Best Local Similarity 100.0%; Pred. No. 0;
FT      Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT      Qy 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGTAGGGCATTCGC 60
FT      2223 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGTAGGGCATTCGC 2282
FT      Db 61 GAGTGTGGGACTTGAACCTGAGCCCGAAGCCCAAGCCCAACCAAGCAAGAGCAGGAC 120
FT      2283 GAGTGTGGGACTTGAACCTGAGCCCGAAGCCCAAGCCCAACCAAGCAAGAGCAGGAC 2342
FT      Qy 121 GACGGCGGGCTCTGGTGTCTTCTGGCTACAAGTACTCTCGGACCTTCAACGGACTCGAC 180
FT      2343 GACGGCGGGCTCTGGTGTCTTCTGGCTACAAGTACTCTCGGACCTTCAACGGACTCGAC 2402
FT      Qy 181 AAGGGGAGCCCGTCAACGCGCGGACGACGCGGCCCTCGAGCAACCAAGGCGCTACGAC 240
FT      2403 AAGGGGAGCCCGTCAACGCGCGGACGACGCGGCCCTCGAGCAACCAAGGCGCTACGAC 2462
FT      Qy 241 CAGCAGCTCAAGCGGGGTGACAAATCGTACTCTCGGTATATACCAACGCGCGGCGGATTT 300
FT      2463 CAGCAGCTCAAGCGGGGTGACAAATCGTACTCTCGGTATATACCAACGCGCGGCGGATTT 2522

```

QY 301 CAGGAGCGTCTGCAAGAGATAGTCTTTTGGGGGCAACCTCGGGGCGAGCAGTCTTCCAG 360  
Db 2523 CAGGAGCGTCTGCAAGAGATAGTCTTTTGGGGGCAACCTCGGGGCGAGCAGTCTTCCAG 2582  
QY 361 GCCAAGACGGGTCTCTCAACCTCTCGGTCTGGTTGAGGAGCGCTAAGACGGTCTCT 420  
Db 2583 GCCAAGACGGGTCTCTCAACCTCTCGGTCTGGTTGAGGAGCGCTAAGACGGTCTCT 2642  
QY 421 GGAAAGAAACGTCCGGTAGAGCAGTCCGCCAAGAGCAGACTCTCTCTCGGCGCATCGGC 480  
Db 2643 GGAAAGAAACGTCCGGTAGAGCAGTCCGCCAAGAGCAGACTCTCTCTCGGCGCATCGGC 2702  
QY 481 AAGACAGGCGCAGCGCCCTTAAAGAGAGACTCAATTTTGGTGCAGACTGGCGACTCAGAG 540  
Db 2703 AAGACAGGCGCAGCGCCCTTAAAGAGAGACTCAATTTTGGTGCAGACTGGCGACTCAGAG 2762  
QY 541 TCAGTCCCGATCCACAACTCTCGGAGAACTCTCCAGCAACCCCGCTCTGTGGGACCT 600  
Db 2763 TCAGTCCCGATCCACAACTCTCGGAGAACTCTCCAGCAACCCCGCTCTGTGGGACCT 2822  
QY 601 ACTACAATGGCTTACGGCGGTGGCGCACCAATGGCAGACAATAAGGAGCGCGCAGCGA 660  
Db 2823 ACTACAATGGCTTACGGCGGTGGCGCACCAATGGCAGACAATAAGGAGCGCGCAGCGA 2882  
QY 661 GTGGGTAAATGCCCTCAGGAAATTTGGCATTTGGATTCCAGATGGCTGGCGACAGATCATC 720  
Db 2883 GTGGGTAAATGCCCTCAGGAAATTTGGCATTTGGCATTTCCAGATGGCTGGCGACAGATCATC 2942  
QY 721 ACCACAGACACCGGACCTTGGGCTTTGCCACCTCAATAACCACTCTTACAGCAAAATC 780  
Db 2943 ACCACAGACACCGGACCTTGGGCTTTGCCACCTCAATAACCACTCTTACAGCAAAATC 3002  
QY 781 TCAGTGTCTTAAACGGGGCGAGCAACGAAACCACTACTTGGCTTACGACACCCCTGG 840  
Db 3003 TCAGTGTCTTAAACGGGGCGAGCAACGAAACCACTACTTGGCTTACGACACCCCTGG 3062  
QY 841 GGGTATTTTGTATTTCAACAGATTCACCTGGACTTTCACGAGTGTGCGAGGACATC 900  
Db 3063 GGGTATTTTGTATTTCAACAGATTTCCACTGCCATTTTCCACAGTGTACTGGCAGCGACTC 3122  
QY 901 ATCAACAAATTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960  
Db 3123 ATCAACAAATTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 3182  
QY 961 GTCAAGGAGGTACGACGAATGATGGCGTCAACCAATCGCTAATAACCTTACCAGCAG 1020  
Db 3183 GTCAAGGAGGTACGACGAATGATGGCGTCAACCAATCGCTAATAACCTTACCAGCAG 3242  
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCCGTAAGTCTCGGCTCTGGGACGAG 1080  
Db 3243 GTTCAAGTCTTCTCGGACTCGGAGTACCAGCTTCCGTAAGTCTCGGCTCTGGGACGAG 3302  
QY 1081 GGTGCTCTCCCTCCGTTCCGGGGGACGTTTCATGATTCGGCAATACGGCTACCTGAGC 1140  
Db 3303 GGTGCTCTCCCTCCGTTCCGGGGGACGTTTCATGATTCGGCAATACGGCTACCTGAGC 3362  
QY 1141 CTCAACAAATGGCAGCAAGCCGTGGGACGTTTCATCTTTTACGCTTGGGAATATTTCCCT 1200  
Db 3363 CTCAACAAATGGCAGCAAGCCGTGGGACGTTTCATCTTTTACGCTTGGGAATATTTCCCT 3422  
QY 1201 TCTCAGATGCTGAGAACGGGCACAACTTTACCTTCAGCTACACTTTGAGGAGTGGCT 1260  
Db 3423 TCTCAGATGCTGAGAACGGGCACAACTTTACCTTCAGCTACACTTTGAGGAGTGGCT 3482  
QY 1261 TTCCACAGCAGCTACGCGACACGCCAGAGCTCGGCGGTGATGATCTCTCTCATCGAC 1320  
Db 3483 TTCCACAGCAGCTACGCGACACGCCAGAGCTCGGCGGTGATGATCTCTCTCATCGAC 3542  
QY 1321 CAATACCTGTATTTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAGGAC 1380  
Db 3543 CAATACCTGTATTTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAGGAC 3602

QY 1381 TTGCTGTTTTAGCCGGTGGTCTCCAGCTGGCATGTCTGTTCCGCCCAAAACCTGGTACT 1440  
Db 3603 TTGCTGTTTTAGCCGGTGGTCTCCAGCTGGCATGTCTGTTCCGCCCAAAACCTGGTACT 3662  
QY 1441 GGACCTCTGTATTCGGCAGCAGCGCTTTCTAAAAAAGAGAGCAACAAACAGCAAT 1500  
Db 3663 GGACCTCTGTATTCGGCAGCAGCGCTTTCTAAAAAAGAGAGCAACAAACAGCAAT 3722  
QY 1501 TTTTACTGGACTGTGTCTTCAAAATATAACCTCAATGGCGTGAATTCATCATCAACCT 1560  
Db 3723 TTTTACTGGACTGTGTCTTCAAAATATAACCTCAATGGCGTGAATTCATCATCAACCT 3782  
QY 1561 GGCACTGTATTCGGCTCTCACAAAGACGAGCAAGTCTTTTCCCATGAGCGGTGTC 1620  
Db 3783 GGCACTGTATTCGGCTCTCACAAAGACGAGCAAGTCTTTTCCCATGAGCGGTGTC 3842  
QY 1621 ATGATTTTGGAAAAAGAGCGCGGAGCTTCAAAACATGCAATTGGACAAATGTCTATGATT 1680  
Db 3843 ATGATTTTGGAAAAAGAGCGCGGAGCTTCAAAACATGCAATTGGACAAATGTCTATGATT 3902  
QY 1681 ACACGAGAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTTGGGACCGTG 1740  
Db 3903 ACACGAGAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTTTGGGACCGTG 3962  
QY 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACCTCGCGAGAGATGTGCATGTATGGGA 1800  
Db 3963 GCAGTCAATTTCCAGAGCAGCAGCAGACCTCGCGAGAGATGTGCATGTATGGGA 4022  
QY 1801 GCATTTACCTGGCATGGTGTGGCAAGTAGAGACGTGTACCTGCAGGGTCCCATTTGGGCC 1860  
Db 4023 GCATTTACCTGGCATGGTGTGGCAAGTAGAGACGTGTACCTGCAGGGTCCCATTTGGGCC 4082  
QY 1861 AAAATTCCTCACAGATGGACATTTTCAACCGTCTCTCTTTATGGCGGCTTTGGACTC 1920  
Db 4083 AAAATTCCTCACAGATGGACATTTTCAACCGTCTCTCTTTATGGCGGCTTTGGACTC 4142  
QY 1921 AGAACCCTCCCTCAGATCTCTCATCAAAACAGCCTGTCTTCGGAATCCTCCGGG 1980  
Db 4143 AGAACCCTCCCTCAGATCTCTCATCAAAACAGCCTGTCTTCGGAATCCTCCGGG 4202  
QY 1981 GAGTTTTAGCTTACAAAGTTTCTTTCATTCACCAATCTCCACAGGACAAGTGA 2040  
Db 4203 GAGTTTTAGCTTACAAAGTTTCTTTCATTCACCAATCTCCACAGGACAAGTGA 4262  
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAG 2100  
Db 4263 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAG 4322  
QY 2101 TACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGACTT 2160  
Db 4323 TACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGACTT 4382  
QY 2161 TATACTGAGCCTCGCCCATTTGGCACCGGTTACCTTTACCGTCCCTGT 2209  
Db 4383 TATACTGAGCCTCGCCCATTTGGCACCGGTTACCTTTACCGTCCCTGT 4431

## RESULT 7

ADE76507

ID ADE76507 standard; DNA; 4718 BP.

XX ADE76507;

AC ADE76507;

XX 29-JAN-2004 (first entry)

XX Adeno-associated virus (AAV) related DNA, SEQ ID No 6.

XX adeno-associated virus; AAV; cytostatic; antipneumatic; antirheumatic;  
XX antithyroid; neuroprotective; antidiabetic; antithyroid;  
XX dermatological; antiinflammatory; gene therapy; vaccine;  
XX hyperproliferative; cancer; psoriasis; autoimmune disease;  
XX rheumatoid arthritis; multiple sclerosis; diabetes;  
XX autoimmune thyroiditis; scleroderma; Crohn's disease; gene; da.

XX Adeno-associated virus 1.  
OS EP1310571-A2.  
XX 14-MAY-2003.  
XX 12-NOV-2002; 2002BP-00257826.  
XX 13-NOV-2001; 2001US-0350607P.  
PR 17-DEC-2001; 2001US-0341117P.  
PR 01-MAY-2002; 2002US-0377066P.  
PR 05-JUN-2002; 2002US-0386675P.  
XX (UIPE-) UNIV PENNSYLVANIA.  
XX Gao G, Wilson JM, Alvira M;  
XX WPI; 2003-450984/43.  
XX  
XX Detecting adeno-associated virus sequences in a sample, useful for e.g.  
PT preventing or treating hyperproliferative or autoimmune diseases,  
PT comprising subjecting a sample having a DNA to amplification via  
PT polymerase chain reaction.  
XX  
XX Claim 14; SEQ ID NO 6; 419pp; English.  
XX  
XX The invention relates to a novel method for detecting adeno-associated  
CC virus (AAV) sequences in a sample, which comprises subjecting a sample  
CC containing a DNA to amplification via a polymerase chain reaction (PCR).  
CC The AAV sequence have the following activities: cytostatic,  
CC antipsoriatic, antirheumatic, antiarthritic, neuroprotective,  
CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV  
CC sequence can be used in gene therapy or as part of a vaccine to treat  
CC disorders. The method is useful in detecting and/or identifying AAV  
CC sequences and isolating novel sequences that are identified. The  
CC sequences may be used e.g. for preventing or treating hyperproliferative  
CC conditions such as cancers and psoriasis, and other autoimmune diseases  
CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune  
CC thyroiditis, scleroderma or Crohn's disease. This polynucleotide sequence  
CC represents an AAV related DNA sequence of the invention.  
XX  
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 2209; DB 10; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCAATTGCG 60  
DB 2223 ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCAATTGCG 2282  
QY 61 GAGTGTGGGACTTGAAACCTTGAGCCCGGAGCCGAAAGCCAAAGCCAAACGAAAGAGGAGGAC 120  
DB 2283 GAGTGTGGGACTTGAAACCTTGAGCCCGGAGCCGAAAGCCAAAGCCAAACGAGGAGGAC 2342  
QY 121 GAGCGCGGGTCTGGTCTTCTGCTCAAGTACTCGGACCCCTTCAACGGACTCGAC 180  
DB 2343 GACGGCGGGGCTTGGTGTCTTCTGGCTCAAGTACTCGGACCCCTTCAACGGACTCGAC 2402  
QY 181 AAGGGGAGCCCGCTCAACCGCGGGGAGCGAGCGGCCCTCGAGCAGCAAGGCTTACGAC 240  
DB 2403 AAGGGGAGCCCGCTCAACCGCGGGGAGCGAGCGGCCCTCGAGCAGCAAGGCTTACGAC 2462  
QY 241 CAGCAGCTCAAGCGGGTGACATCCGTACTTGGGGTATTAACACGCGCGACCGGAGTTT 300  
DB 2463 CAGCAGCTCAAGCGGGTGACATCCGTACTTGGGGTATTAACACGCGCGACCGGAGTTT 2522  
QY 301 CAGGAGCGTCTGCAAGAGATAGTCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 360  
DB 2523 CAGGAGCGTCTGCAAGAGATAGTCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 2582  
QY 361 GCCAAGAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGGCTAAGACGGCTCCT 420

DB 2583 GCCAAGAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTTAAGACGGCTCCT 2642  
QY 421 GGAAAGAAACGTCGGGTAGAGCAGTCGCCACAGAGCCAGACTCTCTCTCGGGCATCGGC 480  
DB 2643 GGAAAGAAACGTCGGGTAGAGCAGTCGCCACAGAGCCAGACTCTCTCTCGGGCATCGGC 2702  
QY 481 AAGACAGGCCACGAGCCCGCTAAAAGAGACTCAATTTTGGTTCAGACTGGGACTCAGAG 540  
DB 2703 AAGACAGGCCACGAGCCCGCTAAAAGAGACTCAATTTTGGTTCAGACTGGGACTCAGAG 2762  
QY 541 TCAGTCCCGGATCCACACCTCTCGGAGAACCTCCAGCAACCCCGCTCTGTGGGACCT 600  
DB 2763 TCAGTCCCGGATCCACACCTCTCGGAGAACCTCCAGCAACCCCGCTCTGTGGGACCT 2822  
QY 601 ACTACAAATGGCTTCAGGGCGGTGGCGCACCAATATGGGAGACAATAACGAAGCGCCGACGGA 660  
DB 2823 ACTACAAATGGCTTCAGGGCGGTGGCGCACCAATATGGGAGACAATAACGAAGCGCCGACGGA 2882  
QY 661 GTGGGTAATGCTCAGGAAATTTGGCAATTCGCAATTCACATGGCTGGGGCAGACAGTCAATC 720  
DB 2883 GTGGGTAATGCTCAGGAAATTTGGCAATTCGCAATTCGCAATGGCTGGGGCAGACAGTCAATC 2942  
QY 721 ACCACGACCGCGACCTGGGGCTTGCCCACTACATAAACCACTCTTACAGCAAAATC 780  
DB 2943 ACCACGACCGCGACCTGGGGCTTGCCCACTACATAAACCACTCTTACAGCAAAATC 3002  
QY 781 TCCAGTGTCTTCAACGGGGCCAGCAACGACCACTACTTCTGGGTTCAGACACCCCTCTGG 840  
DB 3003 TCCAGTGTCTTCAACGGGGCCAGCAACGACCACTACTTCTGGGTTCAGACACCCCTCTGG 3062  
QY 841 GGGTATTTTGAATTTCAACAGATTTCCACTGCCACTTTTACCACGTCGTCGACGAGCACTC 900  
DB 3063 GGGTATTTTGAATTTCAACAGATTTCCACTGCCACTTTTACCACGTCGTCGACGAGCACTC 3122  
QY 901 ATCAACAACATTTGGGGATTCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 960  
DB 3123 ATCAACAACATTTGGGGATTCGGGCCAAGAGACTCAACTTCAAACTCTTCAACATCCAA 3182  
QY 961 GTCAAGGAGGTCAACGACGAATGATGGCGTCACAACCATCGCTAAATAACCTTACGACGACG 1020  
DB 3183 GTCAAGGAGGTCAACGACGAATGATGGCGTCACAACCATCGCTAAATAACCTTACGACGACG 3242  
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCACTTCCGTCTCGGCTCTGGCAACCAAG 1080  
DB 3243 GTTCAAGTCTTCTCGGACTCGGAGTACCACTTCCGTCTCGGCTCTGGCAACCAAG 3302  
QY 1081 GGTGCTCTCCTCCGTTCCGGGGGACGTTTCATGATTTCCGCAATACGGCTACCTGACG 1140  
DB 3303 GGTGCTCTCCTCCGTTCCGGGGGACGTTTCATGATTTCCGCAATACGGCTACCTGACG 3362  
QY 1141 CTCACAAATGGGAGCCCAAGCCGTTGGGACGTTTCATCTCTTTTACTGCTCGGAATATTTCCCT 1200  
DB 3363 CTCACAAATGGGAGCCCAAGCCGTTGGGACGTTTCATCTTTTACTGCTCGGAATATTTCCCT 3422  
QY 1201 TCTCAGATCTGAGAACGGGCAACAACTTTTACCTTTCAGCTTACACCTTTGAGGAAGTGCCT 1260  
DB 3423 TCTCAGATCTGAGAACGGGCAACAACTTTTACCTTTCAGCTTACACCTTTGAGGAAGTGCCT 3482  
QY 1261 TTCCACAGAGCTTACCGGCAAGCGGCTGGACCGGCTGATGATCTCTCATCTGAC 1320  
DB 3483 TTCCACAGAGCTTACCGGCAAGCGGCTGGACCGGCTGATGATCTCTCATCTGAC 3542  
QY 1321 CAATACCTCTATTTACTCTGAACAGACTCAAAATCAGTCCGGAGTCCCAAAACAAGGAC 1380  
DB 3543 CAATACCTCTATTTACTCTGAACAGACTCAAAATCAGTCCGGAGTCCCAAAACAAGGAC 3602  
QY 1381 TTGCTGTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCCAAAACCTGGCTACCT 1440  
DB 3603 TTGCTGTTAGCGGTGGGTCTCCAGCTGGCATGTCTGTTCAGCCCCAAAACCTGGCTACCT 3662  
QY 1441 GGACCTGTTATCGGAGCGGCTTTCTTAAACAAAAACAGACAAACAAACAGCAAT 1500

Db 3663 GGACCTGTTATCGGCAGCGCGGTTCTTAAACAAAAAAGACAGACAAACAGCAACT 3722  
Qy 1501 TTTACTGTGACTGGTGTCTTCAAAATATAAATCAATCAATGGGCGTGAATCAATCAACCTT 1560  
Db 3723 TTTACTGTGACTGGTGTCTTCAAAATATAAATCAATCAATGGGCGTGAATCAATCAACCTT 3782  
Qy 1561 GGCACCTGCTATGGCTTCACAAAAGACGAGCAAGTCTTCTTCCATGAGCGGTGC 1620  
Db 3783 GGCACCTGCTATGGCTTCACAAAAGACGAGCAAGTCTTCTTCCATGAGCGGTGC 3842  
Qy 1621 ATGATTTTTGAAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGACAAATGTCATGATT 1680  
Db 3843 ATGATTTTTGAAAAGAGAGCGCGGAGCTTCAAAACATGCAATGGACAAATGTCATGATT 3902  
Qy 1681 ACAGACGAAGAGAAATTAAGCCACTAACCCTGTGGCCACCGAAGATTGGGACCGTG 1740  
Db 3903 ACAGACGAAGAGAAATTAAGCCACTAACCCTGTGGCCACCGAAGATTGGGACCGTG 3962  
Qy 1741 GCAGTCAATTTCCAGAGCAGCAGACAGACCCCTGCGACCGGAGATGTCATGCGGA 1800  
Db 3963 GCAGTCAATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTCATGCGGA 4022  
Qy 1801 GCATTACCTGGCATGCTGGCAAGATAGAGAGCTGTACCTGCAGGGTCCCAATTTGGGCG 1860  
Db 4023 GCATTACCTGGCATGCTGGCAAGATAGAGAGCTGTACCTGCAGGGTCCCAATTTGGGCG 4082  
Qy 1861 AAAATTCCTCACACAGATGACATTTCAACCGCTCTCTTATGGGCGCTTTGAGCTC 1920  
Db 4083 AAAATTCCTCACACAGATGACATTTCAACCGCTCTCTTATGGGCGCTTTGAGCTC 4142  
Qy 1921 AGAAGCCGCTCTCAGATCCTCATCAAAACAGCGCTGTTCTCGGAATCCTCGGCG 1980  
Db 4143 AGAAGCCGCTCTCAGATCCTCATCAAAACAGCGCTGTTCTCGGAATCCTCGGCG 4202  
Qy 1981 GAGTTTTCAGTACAAAGTTTGTCTTCAATCATCAACCAATACCTCCACAGACAAGTGAGT 2040  
Db 4203 GAGTTTTCAGTACAAAGTTTGTCTTCAATCATCAACCAATACCTCCACAGACAAGTGAGT 4262  
Qy 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCGAAGTGCAG 2100  
Db 4263 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCGAAGTGCAG 4322  
Qy 2101 TACACATCCAAATTTATCAAAATCTGCCAAGTTGATTTTACTGTGGACAAATGACATT 2160  
Db 4323 TACACATCCAAATTTATCAAAATCTGCCAAGTTGATTTTACTGTGGACAAATGACATT 4382  
Qy 2161 TATACTGAGCTCGCCCCATTTGGCACCCGTTACCTTTACCGTCCCTGT 2209  
Db 4383 TATACTGAGCTCGCCCCATTTGGCACCCGTTACCTTTACCGTCCCTGT 4431

RESULT 8

ADL13984  
ID ADL13984 standard; DNA; 4718 BP.  
XX  
AC ADL13984;  
XX

06-MAY-2004 (first entry)

Adeno-associated virus serotype 1 complete DNA.

XX ss; cytostatic; neuroprotective; antiinflammatory; gene therapy;  
XX expression construct; adeno-associated virus;  
XX integration efficiency element; inverted terminal repeat; integration;  
XX chromosome; cancer; lymphoma; leukemia; multiple myeloma; neuroblastoma;  
XX retinoblastoma; inflammatory disease; arthritis;  
XX neurodegenerative disease.

OS Adeno-associated virus 1.

XX WO2003087334-A2.

XX 23-OCT-2003.

PD

XX 09-APR-2003; 2003WO-US011191.  
XX 09-APR-2002; 2002US-0371044P.  
XX (CORR ) CORNELL RES FOUND INC.  
XX Falck-Pedersen ES, Philipott N;  
XX WPI; 2003-833723/77.  
XX New expression construct comprising a nucleic acid sequence encoding an  
XX adeno-associated virus integration efficiency element, useful for  
XX treating cancer e.g. lung cancer or colon cancer or inflammatory disease  
XX e.g. arthritis.  
XX Disclosure; SEQ ID NO 3; 62pp; English.  
XX The invention relates to an expression construct comprising a nucleic  
XX acid sequence encoding an adeno-associated virus integration efficiency  
XX element (AAV IEE), which is devoid of AAV inverted terminal repeats (AAV  
XX ITRs) and site-specifically integrates into a host cell chromosome when  
XX provided to the host cell in conjunction with an AAV Rep protein. The  
XX expression construct can be used as a therapeutic factor for treating a  
XX mammal for a pathologic state which is cancer, including lung cancer,  
XX colon cancer, renal cancer, anal cancer, bile duct cancer, bladder  
XX cancer, bone cancer, brain cancer, spinal chord cancer, breast cancer,  
XX cervical cancer, lymphoma, endometrial cancer, esophageal cancer,  
XX gallbladder cancer, gastrointestinal cancer, laryngeal cancer, leukemia,  
XX liver cancer, multiple myeloma, neuroblastoma, ovarian cancer, pancreatic  
XX cancer, prostatic cancer, retinoblastoma, skin cancer, stomach cancer,  
XX testicular cancer, thymus cancer or thyroid cancer. Other pathologic  
XX state includes inflammatory disease (arthritis), neurodegenerative  
XX disease, a disease of an organ attributed to the presence of increased or  
XX decreased level of a particular gene product(s). This sequence  
XX corresponds to the AAV serotype 1 complete DNA including the IEE  
XX sequence.

XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;

Query Match 100.0%; Score 2209; DB 10; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGCTGGCATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC 60  
Db 2223 ATGGCTGGCATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC 2282

Qy 61 GAGTGGTGGGACTTGAACCTGGAGCCCGGAGCCCAAGCAACAGCAAGCAGGAC 120  
Db 2283 GAGTGGTGGGACTTGAACCTGGAGCCCGGAGCCCAAGCAACAGCAAGCAGGAC 2342

Qy 121 GACGCGCGGGTCTGGTCTTCTGGCTACAGTACCTCGGACCTTCAACGGACTCGAC 180  
Db 2343 GACGCGCGGGTCTGGTCTTCTGGCTACAGTACCTCGGACCTTCAACGGACTCGAC 2402

Qy 181 AAGGGGAGCCCGTCAACCGCGCGGACGCGAGCGGCTCGAGCAGCAAGGCTTACGAC 240  
Db 2403 AAGGGGAGCCCGTCAACCGCGCGGACGCGAGCGGCTCGAGCAGCAAGGCTTACGAC 2462

Qy 241 CAGCAGCTCAAGCGGGTGACAATCCGTACCTGCGGTATAACACGCCGACCCGAGTTT 300  
Db 2463 CAGCAGCTCAAGCGGGTGACAATCCGTACCTGCGGTATAACACGCCGACCCGAGTTT 2522

Qy 301 CAGGAGCGTCTCAAGAAGATACGCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 360  
Db 2523 CAGGAGCGTCTCAAGAAGATACGCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 2582

Qy 361 GCCAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTTAAGACGGTCTCT 420  
Db 2583 GCCAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTTAAGACGGTCTCT 2642

Qy 421 GGAAAGAAACGTCCGGTAGAGCAGTCGCCACAAGACCCAGACTCTCTCTCGGGCATTCGGC 480

Db 2643 GGAAGAAAGCTCCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCTCGGCACTCGGC 2702  
QY 481 AAGACAGCCAGACGCCCTAAAGAGAGACTCAATTTGGTTCAGACTGGCCACTCAGAG 540  
Db 2703 AAGACAGCCAGACGCCCTAAAGAGAGACTCAATTTGGTTCAGACTGGCCACTCAGAG 2762  
QY 541 TCAGTCCCCGATCCACAACCTCTCGAGAACTCCAGCAACCCCGCTGCTGTGGACCT 600  
Db 2763 TCAGTCCCCGATCCACAACCTCTCGAGAACTCCAGCAACCCCGCTGCTGTGGACCT 2822  
QY 601 ACTACAAATGGCTTCAGAGCGGTGGCGCAACAATGGCAGACAATAAAGAAAGCCCGCAGCGA 660  
Db 2823 ACTACAAATGGCTTCAGAGCGGTGGCGCAACAATGGCAGACAATAAAGAAAGCCCGCAGCGA 2882  
QY 661 GTGGGTAAATGGCTTCAGAGAAATGGCAATTCAGATGCTGGCGCAGAGATCATC 720  
Db 2883 GTGGGTAAATGGCTTCAGAGAAATGGCAATTCAGATGCTGGCGCAGAGATCATC 2942  
QY 721 ACCACAGACCCGCACTGGGCTTTGCCCACTTACAATAACCACTCTTACAAGCAAAATC 780  
Db 2943 ACCACAGACCCGCACTGGGCTTTGCCCACTTACAATAACCACTCTTACAAGCAAAATC 3002  
QY 781 TCCAGTGTCTTAAACGGGGCCAGCAACGAACCACTACTTCGGCTTACAGCAACCCCTGG 840  
Db 3003 TCCAGTGTCTTAAACGGGGCCAGCAACGAACCACTACTTCGGCTTACAGCAACCCCTGG 3062  
QY 841 GGGTATTTGATTTCAACAGATTCACATGCTGCTTTCACACAGTGAAGTGGCAGGACTC 900  
Db 3063 GGGTATTTGATTTCAACAGATTCACATGCTGCTTTCACACAGTGAAGTGGCAGGACTC 3122  
QY 901 ATCAACAAATTTGGGATTCGGGCCAAGAGACTCAACTTCAAACTTTCAACATCCAA 960  
Db 3123 ATCAACAAATTTGGGATTCGGGCCAAGAGACTCAACTTCAAACTTTCAACATCCAA 3182  
QY 961 GTCAAGGAGTCAACAGCAATGATGGCGTCAACCAATCGCTTAATCACTTACCGACG 1020  
Db 3183 GTCAAGGAGTCAACAGCAATGATGGCGTCAACCAATCGCTTAATCACTTACCGACG 3242  
QY 1021 GTTCAAGTCTTTCGAGCTCGAGTACAGACTTCGGTACGCTCGGCTCTGGCAGCAG 1080  
Db 3243 GTTCAAGTCTTTCGAGCTCGAGTACAGACTTCGGTACGCTCGGCTCTGGCAGCAG 3302  
QY 1081 GGTGCTCTCTCCGTTCCCGCGGAGCTTTCATGATTCGCAATACGCTACCTGAGC 1140  
Db 3303 GGTGCTCTCTCCGTTCCCGCGGAGCTTTCATGATTCGCAATACGCTACCTGAGC 3362  
QY 1141 CTCAACAATGGCAGCAAGCCGTGGGACGTTCACTCTTTTACTGCTTGAATATTTCCCT 1200  
Db 3363 CTCAACAATGGCAGCAAGCCGTGGGACGTTCACTCTTTTACTGCTTGAATATTTCCCT 3422  
QY 1201 TCTCAGATGCTGAGAACGGGCAACAATTTACTTTCAGCTACACCTTTGAGGAGTGCCT 1260  
Db 3423 TCTCAGATGCTGAGAACGGGCAACAATTTACTTTCAGCTACACCTTTGAGGAGTGCCT 3482  
QY 1261 TTCCACAGAGCTTACCGCACAGCAGCTCGACCGGCTGATGATCTCTCATCGAC 1320  
Db 3483 TTCCACAGAGCTTACCGCACAGCAGCTCGACCGGCTGATGATCTCTCATCGAC 3542  
QY 1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAGGAC 1380  
Db 3543 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAAGGAC 3602  
QY 1381 TTGCTGTTAGCGTGGGCTTCAGCTGGGATGTCTGTTGAGCCCAAAACTGGCTACCT 1440  
Db 3603 TTGCTGTTAGCGTGGGCTTCAGCTGGGATGTCTGTTGAGCCCAAAACTGGCTACCT 3662  
QY 1441 GGAACCTGTATTCCGCGAGCGCGGTTCTTAAACAAACACACACACACACACAAAT 1500  
Db 3663 GGAACCTGTATTCCGCGAGCGCGGTTCTTAAACAAACACACACACACACAAAT 3722  
QY 1501 TTTTACCTGGACTGGTGTCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCCCT 1560

Db 3723 TTTTACCTGGACTGGTGTCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCCCT 3782  
QY 1561 GGCACCTGTATGGCTCTCACAAAAGACGACGAAGCAAAAGTTCTTTCCATGAGCGGTGTC 1620  
Db 3783 GGCACCTGTATGGCTCTCACAAAAGACGACGAAGCAAAAGTTCTTTCCATGAGCGGTGTC 3842  
QY 1621 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCATTGGACAAATGTCTATGTT 1680  
Db 3843 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAAAACACTGCATTGGACAAATGTCTATGTT 3902  
QY 1681 ACAGCAGAGAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTTGGACCCGTG 1740  
Db 3903 ACAGCAGAGAGAGAAATTAAGCCCACTAACCTGTGGCCACCGAAGATTTTGGACCCGTG 3962  
QY 1741 GCAGTCAATTTTCCAGAGCAGCAGACAGACCCCTGCGACCGGAGATGTGCTATGGGA 1800  
Db 3963 GCAGTCAATTTTCCAGAGCAGCAGACAGACCCCTGCGACCGGAGATGTGCTATGGGA 4022  
QY 1801 GCATTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGCTCCCATTTGGGCC 1860  
Db 4023 GCATTACCTGGCATGGTGTGGCAAGATAGAGACGTGTACCTGCAGGCTCCCATTTGGGCC 4082  
QY 1861 AAAATTCCTTCACACAGATGGACACTTTTCAACCGCTCTCTCTTATGGGCGGCTTTGGACTC 1920  
Db 4083 AAAATTCCTTCACACAGATGGACACTTTTCAACCGCTCTCTCTTATGGGCGGCTTTGGACTC 4142  
QY 1921 AAGAACCCGCTCTCTCAGATCTCTCAATAAACAAGCCCTGTTCTCCTGCAATCTCCGGCG 1980  
Db 4143 AAGAACCCGCTCTCTCAGATCTCTCAATAAACAAGCCCTGTTCTCCTGCAATCTCCGGCG 4202  
QY 1981 GAGTTTTCAGCTTACAAAGTTTCTTCAATTCATCACCAATATCTCACAGGACAAGTGACT 2040  
Db 4203 GAGTTTTCAGCTTACAAAGTTTCTTCAATTCATCACCAATATCTCACAGGACAAGTGACT 4262  
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCGCTGGAATCCCGAAGTGCAG 2100  
Db 4263 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCGCTGGAATCCCGAAGTGCAG 4322  
QY 2101 TACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGACTT 2160  
Db 4323 TACACATCCAAATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGACTT 4382  
QY 2161 TATACTGAGCTTCGCCCAATTTGGCACCCGTTACCTTACCTGCTCCCTGT 2209  
Db 4383 TATACTGAGCTTCGCCCAATTTGGCACCCGTTACCTTACCTGCTCCCTGT 4431

RESULT 9  
ADG39758  
ID ADG39758 standard; DNA; 4718 BP.  
XX  
AC ADG39758;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
AAV-1 genomic DNA sequence SEQ ID NO:20.  
XX  
parvovirus; rep; cap; DNA binding domain; capsid interacting domain;  
KW recombinant hybrid parvovirus particle;  
KW recombinant adeno-associated virus; rAAV; AAV; vaccine; gene therapy;  
gene; ds.  
XX  
Adeno-associated virus 1.  
XX  
WO2003104392-A2.  
XX  
18-DEC-2003.  
XX  
02-DEC-2002; 2002WO-US038423.  
XX  
18-DEC-2001; 2001US-0341919P.  
XX  
(UYN-) UNIV NORTH CAROLINA.



XX Samulski RJ, Rabinowitz JE;  
XX WPI; 2004-062324/06.  
XX  
XX New polynucleotides comprising parvovirus rep coding sequences and  
XX parvovirus cap coding sequences, useful in producing higher stocks of  
XX hybrid parvovirus vectors for delivering therapeutic nucleic acids to a  
XX subject.  
XX  
XX Disclosure; SEQ ID NO 20; 115bp; English.  
XX  
XX The present invention describes a polynucleotide (I), comprising  
XX parvovirus rep coding sequences and parvovirus cap coding sequences. The  
XX rep coding sequences encodes a DNA binding domain from a first  
XX parvovirus, and a capsid interacting domain from a parvovirus different  
XX from the first parvovirus. The cap coding sequence comprises sequences  
XX from the different parvovirus. Also described: (1) a vector comprising  
XX (I); (2) a cell comprising (I), or parvovirus rep coding sequences and  
XX parvovirus cap coding sequences, where the rep coding sequences encode a  
XX DNA binding domain from a first parvovirus and a capsid interacting  
XX domain from a parvovirus different from the first parvovirus, the cap  
XX coding sequences comprise sequences from the different parvovirus, and  
XX the rep coding sequences are stably integrated into the genome of the  
XX cell; and (3) producing a recombinant hybrid parvovirus particle or adeno  
XX -associated virus (rAAV) particle. (I) can be used in vaccines, and in  
XX gene therapy. The polynucleotide (I) can be used in producing higher  
XX stocks of hybrid parvoviruses or parvovirus vectors, which may be used in  
XX the delivery of nucleic acids having biological effect to treat or  
XX ameliorate the symptoms associated with any disorder related to gene  
XX expression. The polynucleotide may be used to produce a parvovirus vector  
XX to express an immunogenic polypeptide in a subject, e.g. for vaccination.  
XX The parvovirus vector may also be used to provide an antisense nucleic  
XX acid to a cell in vitro or in vivo, or in diagnostic and screening  
XX methods. The present sequence is used in the exemplification of the  
XX present invention.  
XX  
XX Sequence 4718 BP; 1121 A; 1393 C; 1273 G; 931 T; 0 U; 0 Other;  
XX  
XX Query Match 100.0%; Score 2209; DB 12; Length 4718;  
XX Best Local Similarity 100.0%; Pred. No. 0;  
XX Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1 ATGGCTGCCGATGGTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60  
XX 2223 ATGGCTGCCGATGGTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 2282  
XX  
XX 61 GAGTGGTGGACCTTGAAACCTGGAGCCCGAAGCCCAAGCCCAAGCCCAAGCAGGAG 120  
XX 2283 GAGTGGTGGACCTTGAAACCTGGAGCCCGAAGCCCAAGCCCAAGCAGGAGGAG 2342  
XX  
XX 121 GACGGCCGGGCTCGTGGTCTTCGCTCAAGTACCTCGGACCTTCAACCGGACTCGAC 180  
XX 2343 GACGGCCGGGCTCGTGGTCTTCGCTCAAGTACCTCGGACCTTCAACCGGACTCGAC 2402  
XX  
XX 181 AAGGGGAGCCCTCAACGGCGGACGAGCGGCGCTCGAGCAGCAGAGGCTTAGCAG 240  
XX 2403 AAGGGGAGCCCTCAACGGCGGACGAGCGGCGCTCGAGCAGCAGAGGCTTAGCAG 2462  
XX  
XX 241 CAGCAGCTCAAGCGGGTGACATCCGTACTCGGGTATAACCAAGCCGACGCGGAGTTT 300  
XX 2463 CAGCAGCTCAAGCGGGTGACATCCGTACTCGGGTATAACCAAGCCGACGCGGAGTTT 2522  
XX  
XX 301 CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGGCAACCTCGGGGAGCAGCTTCCAG 360  
XX 2523 CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGGCAACCTCGGGGAGCAGCTTCCAG 2582  
XX  
XX 361 GCCAAGAGCGGGTCTCGAAGCTCTCGTCTGGTTGAGGAGCGGCTAAGCGGCTCCT 420  
XX 2583 GCCAAGAGCGGGTCTCGAAGCTCTCGTCTGGTTGAGGAGCGGCTAAGCGGCTCCT 2642  
XX  
XX 421 GGAAAGAAACGTCGGGTAGAGCAGTGCACAGAGCAGACCTCTCTCGGGCATCGGC 480  
XX  
XX

Db 2643 GGAAAGAAACGTCGGGTAGAGCAGTGCACAGAGCCAGACTCTCTCTCGGGCATCGGC 2702  
Qy 481 AAGACAGCCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTTCAGACTGGGAGCTCAGAG 540  
Db 2703 AAGACAGCCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTTCAGACTGGGAGCTCAGAG 2762  
Qy 541 TCAGTCCCGGATCCACAACTCTCGGAGAACTCCAGCAACCCCGCTGCTGTGGGACT 600  
Db 2763 TCAGTCCCGGATCCACAACTCTCGGAGAACTCCAGCAACCCCGCTGCTGTGGGACT 2822  
Qy 601 ACTCAATGGCTTCAGGCGGTGGGCGACCAATGGCAGACAATAACGAAGGCGCCGACGA 660  
Db 2823 ACTCAATGGCTTCAGGCGGTGGGCGACCAATGGCAGACAATAACGAAGGCGCCGACGA 2882  
Qy 661 GTGGTAAATGCTCAGGAAATTTGGCAATTTCCATGCTGGGCGAGAGTCAATC 720  
Db 2883 GTGGGTAATGCTCAGGAAATTTGGCAATTTGGCAATTTCCATGCTGGGCGAGAGTCAATC 2942  
Qy 721 ACCACGACCCCGCAGCTGGGCGCTTGGCCACCTTACATAACCACTCTTCAAGCAATC 780  
Db 2943 ACCACGACCCCGCAGCTGGGCGCTTGGCCACCTTACATAACCACTCTTCAAGCAATC 3002  
Qy 781 TCCAGTGTCTTCAACGGGGCGCAGCAACGACCACTACTTTCGGCTACAGCACCCCTGG 840  
Db 3003 TCCAGTGTCTTCAACGGGGCGCAGCAACGACCACTACTTTCGGCTACAGCACCCCTGG 3062  
Qy 841 GGGTATTTTGAATTTCAACAGATTTCACTGGCACTTTTCCACACGTGACTGGCAGGACTC 900  
Db 3063 GGGTATTTTGAATTTCAACAGATTTCACTGGCACTTTTCCACACGTGACTGGCAGGACTC 3122  
Qy 901 ATCAACAACAAATTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCAA 960  
Db 3123 ATCAACAACAAATTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCAA 3182  
Qy 961 GTCAAGGAGGTACACGACGAATGATGGCGTCAACCACTCGCTTAATTAACCTTACCGACAG 1020  
Db 3183 GTCAAGGAGGTACACGACGAATGATGGCGTCAACCACTCGCTTAATTAACCTTACCGACAG 3242  
Qy 1021 GTTCAAGTCTTTCGGACTCGGAGTACAGCTTTCGGTACGCTCTCGGCTTCGGGACCCAG 1080  
Db 3243 GTTCAAGTCTTTCGGACTCGGAGTACAGCTTTCGGTACGCTCTCGGCTTCGGGACCCAG 3302  
Qy 1081 GGTCCCTCCCTCCGCTCCGCGGACGCTTTCATGATTCGCAATACGGCTACCTGACG 1140  
Db 3303 GGTCCCTCCCTCCGCTCCGCGGACGCTTTCATGATTCGCAATACGGCTACCTGACG 3362  
Qy 1141 CTCACAATGGCAGCAAGCCGCTGGGACGTTTCATCTTTTACTGCTCGAAATATTTCCCT 1200  
Db 3363 CTCACAATGGCAGCAAGCCGCTGGGACGTTTCATCTTTTACTGCTCGAAATATTTCCCT 3422  
Qy 1201 TCTCAGATGCTGAGAAACGGGCAACAACTTTTACCTTCAGTACCTTTTGAAGGAGTGCCT 1260  
Db 3423 TCTCAGATGCTGAGAAACGGGCAACAACTTTTACCTTCAGTACCTTTTGAAGGAGTGCCT 3482  
Qy 1261 TTCCACAGCAGCTACCGCAGCAGCAGCTGGACCGGCTGATGAATCTCTCTCATCGAC 1320  
Db 3483 TTCCACAGCAGCTACCGCAGCAGCAGCTGGACCGGCTGATGAATCTCTCTCATCGAC 3542  
Qy 1321 CAATACCTGTATTACTCTGAACAGAACTCAAAATTCAGTTCGGGAGTGCCTCAAAACAGGAC 1380  
Db 3543 CAATACCTGTATTACTCTGAACAGAACTCAAAATTCAGTTCGGGAGTGCCTCAAAACAGGAC 3602  
Qy 1381 TTGCTGTTTTCAGCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAACTGGGTACCT 1440  
Db 3603 TTGCTGTTTTCAGCGTGGGTCTCCAGCTGGCATGTCTGTTTCAGCCCAAACTGGGTACCT 3662  
Qy 1441 GGACCTGTATTTCGGCAGCAGCGCTTTCTAAAAAAGAGAGAGTGCCTCAAAACAGGAC 1500  
Db 3663 GGACCTGTATTTCGGCAGCAGCGCTTTCTAAAAAAGAGAGAGTGCCTCAAAACAGGAC 3722  
Qy 1501 TTTTACCTGGACTGGTCTTCAAAATATAAATCAATTCAGTGGGCGTGAATTCATCAACCT 1560  
Db 3723 TTTTACCTGGACTGGTCTTCAAAATATAAATCAATTCAGTGGGCGTGAATTCATCAACCT 3782



661 GTGGGTAATGCTCAGGAAATTTGGCAATTCGCAATTCGATGCTGGGCGACAGAGTCATC 720  
Db |||||  
5281 GTGGGTAATGCTCAGGAAATTTGGCAATTCGCAATTCGATGCTGGGCGACAGAGTCATC 5340  
Qy |||||  
721 ACCACGACGACCGGACCTTGGGCTTTGGCCACCTTCAATTAACCACTCTTACAAGCAAAATC 780  
Db |||||  
5341 ACCACGACGACCGGACCTTGGGCTTTGGCCACCTTCAATTAACCACTCTTACAAGCAAAATC 5400  
Qy |||||  
781 TCAGTGTCTCAACGGGGGCCAGCAACGACCACTACTTCCGCTACAGCAACCCCTGG 840  
Db |||||  
5401 TCAGTGTCTCAACGGGGGCCAGCAACGACCACTACTTCCGCTACAGCAACCCCTGG 5460  
Qy |||||  
841 GGGTATTTTGAATTTCAACAGATTTCCACTCGCACTTTTACACGAGTGAGTGGGAGGACTC 900  
Db |||||  
5461 GGGTATTTTGAATTTCAACAGATTTCCACTCGCACTTTTACACGAGTGAGTGGGAGGACTC 5520  
Qy |||||  
901 ATCAACAACAATTTGGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTTCAACATCCAA 960  
Db |||||  
5521 ATCAACAACAATTTGGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTTCAACATCCAA 5580  
Qy |||||  
961 GTCAGGAGGTACGACGAAATGATGGCGTACAAACATCGCTAATAACCTTACGAGCAG 1020  
Db |||||  
5581 GTCAGGAGGTACGACGAAATGATGGCGTACAAACATCGCTAATAACCTTACGAGCAG 5640  
Qy |||||  
1021 GTTCAAGTCTTCTCGGACTCGGAGTACCACTTCCGTACGTCTCGGCTCTGGCACCAG 1080  
Db |||||  
5641 GTTCAAGTCTTCTCGGACTCGGAGTACCACTTCCGTACGTCTCGGCTCTGGCACCAG 5700  
Qy |||||  
1081 GGTGCTCTCCCTCGTTCCTCGGCGGACGTGTTCAATTCGCAATACCGCTACCTTGAGC 1140  
Db |||||  
5701 GGTGCTCTCCCTCGTTCCTCGGCGGACGTGTTCAATTCGCAATACCGCTACCTTGAGC 5760  
Qy |||||  
1141 CTCAACAATTTGGGCGGCAAGCGTGGGACGTTTCAATTCCTTTTACGCTGGAAATTTTCCCT 1200  
Db |||||  
5761 CTCAACAATTTGGGCGGCAAGCGTGGGACGTTTCAATTCCTTTTACGCTGGAAATTTTCCCT 5820  
Qy |||||  
1201 TCTCAGATCTGAGACGGGCAACACTTTTACCTTACGCTACACCTTTGAGGAGTGCCCT 1260  
Db |||||  
5821 TCTCAGATCTGAGACGGGCAACACTTTTACCTTACGCTACACCTTTGAGGAGTGCCCT 5880  
Qy |||||  
1261 TTCCACGACGCTTACGCGCACGACGAGCTGGACCGGCTGATGATCTCTCATCGAC 1320  
Db |||||  
5881 TTCCACGACGCTTACGCGCACGACGAGCTGGACCGGCTGATGATCTCTCATCGAC 5940  
Qy |||||  
1321 CAATACCTGTAATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAGGAC 1380  
Db |||||  
5941 CAATACCTGTAATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAGGAC 6000  
Qy |||||  
1381 TTGCTGTTAGCCGTGGGTCTCCAGCTGGCATGCTGTTTCAGCCCAAAACTGGCTACCT 1440  
Db |||||  
6001 TTGCTGTTAGCCGTGGGTCTCCAGCTGGCATGCTGTTTCAGCCCAAAACTGGCTACCT 6060  
Qy |||||  
1441 GGACCTCTGTTATCGGCGACGCGCTTTCTTAAACCAAAACAGACCAACCAACAGCAAT 1500  
Db |||||  
6061 GGACCTCTGTTATCGGCGACGCGCTTTCTTAAACCAAAACAGACCAACCAACAGCAAT 6120  
Qy |||||  
1501 TTTACCTGGACTGGTGTCTTAAATAATAACCTCAATGGGCGTGAATCCATCATCAACCT 1560  
Db |||||  
6121 TTTACCTGGACTGGTGTCTTAAATAATAACCTCAATGGGCGTGAATCCATCATCAACCT 6180  
Qy |||||  
1561 GGCACTGTATGGCTTACAAAGACGAGCAAGTCTTCCCATGAGCGGTGC 1620  
Db |||||  
6181 GGCACTGTATGGCTTACAAAGACGAGCAAGTCTTCCCATGAGCGGTGC 6240  
Qy |||||  
1621 ATGATTTTGGAAAAGAGCGCGGAGCTTTCAACACATGCAATTTGGACAATGTCATGATT 1680  
Db |||||  
6241 ATGATTTTGGAAAAGAGCGCGGAGCTTTCAACACATGCAATTTGGACAATGTCATGATT 6300  
Qy |||||  
1681 ACAGACGAAGAGAAATTAAGCCACTTAACCTGTGGCCACCAAGAAAGATTTGGGACCGTG 1740  
Db |||||  
6301 ACAGACGAAGAGAAATTAAGCCACTTAACCTGTGGCCACCAAGAAAGATTTGGGACCGTG 6360  
Qy |||||  
1741 GCAGTCAATTTCCAGAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCGATGCTATGGGA 1800

6361 GCAGTCAATTTCCAGAGCAGCAGCAGACAGCCTTCGGACCGAGATGTGATGCTATGGGA 6420  
Qy |||||  
1801 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTTGACCTGCGAGGTTCCCATTTGGGCC 1860  
Db |||||  
6421 GCATTTACCTGGCATGGTGTGGCAAGATAGAGACGTTGACCTGCGAGGTTCCCATTTGGGCC 6480  
Qy |||||  
1861 AAAATTTCTCACACAGATGGACATTTTCAACCGGCTCTCTTATGGGCGGCTTTGGACTC 1920  
Db |||||  
6481 AAAATTTCTCACACAGATGGACATTTTCAACCGGCTCTCTTATGGGCGGCTTTGGACTC 6540  
Qy |||||  
1921 AAGAACCCGCTCTCTCAGATCCTCATCAAAAACACGCTGTTCTTCCGAAATCCTCCGGCG 1980  
Db |||||  
6541 AAGAACCCGCTCTCTCAGATCCTCATCAAAAACACGCTGTTCTTCCGAAATCCTCCGGCG 6600  
Qy |||||  
1981 GAGTTTTCAGCTACAAAGTTTGTCTTATCATACCCCAATATCTCACAGGCAAGTGAAT 2040  
Db |||||  
6601 GAGTTTTCAGCTACAAAGTTTGTCTTATCATACCCCAATATCTCACAGGCAAGTGAAT 6660  
Qy |||||  
2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGAG 2100  
Db |||||  
6661 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGAG 6720  
Qy |||||  
2101 TACACATCCAAATTTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGCACTT 2160  
Db |||||  
6721 TACACATCCAAATTTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGCACTT 6780  
Qy |||||  
2161 TATACGAGCTCGCCCAATTTGGCACCCGTTACCTTACCTACCCGTCCTGT 2209  
Db |||||  
6781 TATACGAGCTCGCCCAATTTGGCACCCGTTACCTTACCTACCCGTCCTGT 6829

## RESULT 11

ADZ26928

ID ADZ26928 standard, DNA; 2211 BP.

XX AC ADZ26928;

XX XX 30-JUN-2005 (first entry)

XX DE Adeno-associated virus DNA SEQ ID NO 78.

XX KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;  
XX KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;  
XX KW bacterial infection; cancer; ulcerative colitis; antineumatic;  
XX KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;  
XX KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;  
XX KW virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.

XX OS Adeno-associated virus.

XX PN WO2005033321-A2.

XX PD 14-APR-2005.

XX PF 30-SEP-2004; 2004WO-US028817.

XX XX 30-SEP-2003; 2003US-0508226P.

XX PR 29-APR-2004; 2004US-0566546P.

XX XX (UYPE-) UNIV PENNSYLVANIA.

XX XX Wilson JM, Gao G, Alvira MR, Vandenberghe LH;

XX DR WPI; 2005-285437/29.

XX PT New adeno-associated virus (AAV) clade comprising at least three AAV  
XX PT members, useful for preventing and/or treating arthritis, multiple  
XX PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial  
XX PT infection and cancer.

XX PS Claim 19; SEQ ID NO 78; 569pp; English.

CC The invention relates to an adeno-associated virus (AAV) clade comprising  
CC at least three AAV members, where each member of the AAV clade is  
CC phylogenetically related as determined using a Neighbor-Joining heuristic  
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson  
CC correction distance measurement of no more than 0.05. The methods and  
CC compositions of the present invention are useful for the prevention  
CC and/or treatment of rheumatoid arthritis, multiple sclerosis, Crohn's  
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's  
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative  
CC colitis. The present sequence represents an adeno-associated virus DNA.  
XX

SQ Sequence 2211 BP; 575 A; 654 C; 541 G; 441 T; 0 U; 0 Other;

Query Match 99.4%; Score 2196.2; DB 14; Length 2211;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 2201; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY	1	ATGGCTGCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCAATCGC	60
DB	1	ATGGCTGCGGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGCAATCGC	60
QY	61	GAGTGTGGGACTTGAACCTCGAGCCCGGAGCCCAAGCCCAAGCCCAAGCAGGAGGAC	120
DB	61	GAGTGTGGGACTTGAACCTCGAGCCCGGAGCCCGGAGCCCAAGCCCAAGCAGGAGGAC	120
QY	121	GACGGCCGGGCTCTGGTGTCTTCTGGCTACAAGTACTCGGACCTTCAACGAGCTCGAC	180
DB	121	GACGGCCGGGCTCTGGTGTCTTCTGGCTACAAGTACTCGGACCTTCAACGAGCTCGAC	180
QY	181	AAGGGGAGCCCTCAACGGCGGAGCGAGCGGCTCGAGCAACGAGGCTACGAC	240
DB	181	AAGGGGAGCCCTCAACGGCGGAGCGAGCGGCTCGAGCAACGAGGCTACGAC	240
QY	241	CAGCAGCTCAAGCGGGTGACATCCGTACTCGGTATTAACACCGGAGCGGATTT	300
DB	241	CAGCAGCTCAAGCGGGTGACATCCGTACTCGGTATTAACACCGGAGCGGATTT	300
QY	301	CAGGAGCTCTCAAGAAGATAGCTCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG	360
DB	301	CAGGAGCTCTCAAGAAGATAGCTCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG	360
QY	361	GCCAAAGCGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGCGGCTCCT	420
DB	361	GCCAAAGCGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCTAAGCGGCTCCT	420
QY	421	GGAAGAAAGCTCCGGTAGAGAGTCCGCAACAGAGCCAGACTCTCTCGGGCATCGGC	480
DB	421	GGAAGAAAGCTCCGGTAGAGAGTCCGCAACAGAGCCAGACTCTCTCGGGCATCGGC	480
QY	481	AAGACAGGCGCAGCGCTCAAAAGAGACTCAATTTTGGTCCAGACTGGCGACTCAGAG	540
DB	481	AAGACAGGCGCAGCGCTCAAAAGAGACTCAATTTTGGTCCAGACTGGCGACTCAGAG	540
QY	541	TCAGTCCCCGATCCAAACCTCTCGGAGAACTCCAGCAACCCCGCTGTGTGGACCT	600
DB	541	TCAGTCCCCGATCCAAACCTCTCGGAGAACTCCAGCAACCCCGCTGTGTGGACCT	600
QY	601	ACTACAAATGGCTTACGGCGTGGGCAACCAATGGCAGACAAATACGAGGCGCGGCGGA	660
DB	601	ACTACAAATGGCTTACGGCGTGGGCAACCAATGGCAGACAAATACGAGGCGCGGCGGA	660
QY	661	GTGGGTAAATGCTTCAGGAATTTGGCATTTGGATTTCACATGGCTGGGCGACAGATCATC	720
DB	661	GTGGGTAAATGCTTCAGGAATTTGGCATTTGGCATTTGGCATTTGGCGACAGATCATC	720
QY	721	ACCACAGCACCGGCTTGGGCTTGGCCACCTTACAAATACCACTCTTCAAGCAAAATC	780
DB	721	ACCACAGCACCGGCTTGGGCTTGGCCACCTTACAAATACCACTCTTCAAGCAAAATC	780
QY	781	TCCAGTGTCTTCAACGGGGGCGAGCAACGACCACTACTTTCGGCTACAGCACCCCTCGG	840
DB	781	TCCAGTGTCTTCAACGGGGGCGAGCAACGACCACTACTTTCGGCTACAGCACCCCTCGG	840

QY	841	GGGTATTTTGATTTCAACAGATTCCACTGCGCACTTTTCCACACGCTGACTGGCAGGACTC	900
DB	841	GGGTATTTTGATTTCAACAGATTCCACTGCGCACTTTTCCACACGCTGACTGGCAGGACTC	900
QY	901	ATCAACAACNATGGGGATTCGGGCCAGAGACTCAACTTCAAACTCTTCAACATCCAA	960
DB	901	ATCAACAACNATGGGGATTCGGGCCAGAGACTCAACTTCAAACTCTTCAACATCCAA	960
QY	961	GTCAAGGAGGTCAACGCAATGATGGCGTCAACCACTCGCTTAATACCTTACCAGCAG	1020
DB	961	GTCAAGGAGGTCAACGCAATGATGGCGTCAACCACTCGCTTAATACCTTACCAGCAG	1020
QY	1021	GTTCAGTCTTCTCGGACTTCGAGTACAGACTTCCGTAAGTCTCGGTCTCGGCTCGGAC	1080
DB	1021	GTTCAGTCTTCTCGGACTTCGAGTACAGACTTCCGTAAGTCTCGGTCTCGGCTCGGAC	1080
QY	1081	GGCTGCTCCCTCCGTTCCGGCGGACGTTTCATGATTCGCAATACGCTTACCTGAGC	1140
DB	1081	GGCTGCTCCCTCCGTTCCGGCGGACGTTTCATGATTCGCAATACGCTTACCTGAGC	1140
QY	1141	CTCAACAATGGCAGCCGCTGGGAGCGTTTCTTCTTACTGCTGGAATATTTTCCCT	1200
DB	1141	CTCAACAATGGCAGCCGCTGGGAGCGTTTCTTCTTACTGCTGGAATATTTTCCCT	1200
QY	1201	TCTCAGATGCTCAGAAAGCGGCAACCTTTTACTCTTCTTCTTCTTCTTCTTCTTCTT	1260
DB	1201	TCTCAGATGCTCAGAAAGCGGCAACCTTTTACTCTTCTTCTTCTTCTTCTTCTTCTT	1260
QY	1261	TTTCCAGGAGCTACGGCGACAGCGGCTTCCGCAATACGCTTCTTCTTCTTCTTCTTCT	1320
DB	1261	TTTCCAGGAGCTACGGCGACAGCGGCTTCCGCAATACGCTTCTTCTTCTTCTTCTTCT	1320
QY	1321	CAATACCTCTATTTACTCTGAAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAGG	1380
DB	1321	CAATACCTCTATTTACTCTGAAACAGAACTCAAAATCAGTCCGGAAGTCCCAAAACAGG	1380
QY	1381	TTTGTGTTTGTAGCGGTCTCAGCTGGCATGTCTTTCAGCCCAAACTTCTTCTTCTTCT	1440
DB	1381	TTTGTGTTTGTAGCGGTCTCAGCTGGCATGTCTTTCAGCCCAAACTTCTTCTTCTTCT	1440
QY	1441	GGACCTGTTTATCGGCGAGCGGCTTCTTAAACAAACAGACAGACAAACAGCAAT	1500
DB	1441	GGACCTGTTTATCGGCGAGCGGCTTCTTAAACAAACAGACAGACAAACAGCAAT	1500
QY	1501	TTTACTCTGACTGTGCTTCAAAATATAACCTCAATGGCGGTGAATTCATCAACCTCT	1560
DB	1501	TTTACTCTGACTGTGCTTCAAAATATAACCTCAATGGCGGTGAATTCATCAACCTCT	1560
QY	1561	GGCAGTGTATGGCTTCAACAAAGACGAGCAAGTCTTCTTCTTCTTCTTCTTCTTCTT	1620
DB	1561	GGCAGTGTATGGCTTCAACAAAGACGAGCAAGTCTTCTTCTTCTTCTTCTTCTTCTT	1620
QY	1621	ATGATTTTGGAAAGAGAGCGCGGCTTCAACACTTGGCATTTGGCAATGTTCATGATT	1680
DB	1621	ATGATTTTGGAAAGAGAGCGCGGCTTCAACACTTGGCATTTGGCAATGTTCATGATT	1680
QY	1681	ACAGCAGGAGAGGAAATTAAGCCACTTAACCTTCAATGGCGGTGAATTCATGATT	1740
DB	1681	ACAGCAGGAGAGGAAATTAAGCCACTTAACCTTCAATGGCGGTGAATTCATGATT	1740
QY	1741	GCAGTCAATTTCCAGAGCAGCAGACAGACCTTCCGACCGGAGATGTGATGTATGGGA	1800
DB	1741	GCAGTCAATTTCCAGAGCAGCAGACAGACCTTCCGACCGGAGATGTGATGTATGGGA	1800
QY	1801	GCATTAACCTGGCATGTGTGGCAAGATAGAGAGCTGTACTGCGAGGTTCCCATTTGGGC	1860
DB	1801	GCATTAACCTGGCATGTGTGGCAAGATAGAGAGCTGTACTGCGAGGTTCCCATTTGGGC	1860
QY	1861	AAATTTCTTCAACAGATGGACATTTTCAACCGGTCTTCTTATTTGGCGGCTTGGACTC	1920
DB	1861	AAATTTCTTCAACAGATGGACATTTTCAACCGGTCTTCTTATTTGGCGGCTTGGACTC	1920
QY	1921	AAGAACCCGCTCTCTCAGATCTCTATCAAAAAACAGCTGTCTTCTTCTTCTTCTTCTT	1980

Db 1921 AAGAACCGCCCTCTCAGATCTCATCAAAACACGCGCTGTTCTCGGAATCTCTCGGCG 1980  
QY 1981 GAGTTTTCAGCTACAAAGTTTGGTTCATTCATCAACCCAAATCTCCACAGGACAAAGTGAGT 2040  
Db 1981 GAGTTTTCAGCTACAAAGTTTGGTTCATTCATCAACCCAAATCTCCACAGGACAAAGTGAGT 2040  
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAG 2100  
Db 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAG 2100  
QY 2101 TACACATCAATATATGCAAAATCTGCCAACGTTGATTTTACCTGTGGACAAACAAATGGACTT 2160  
Db 2101 TACACATCAATATATGCAAAATCTGCCAACGTTGATTTTACCTGTGGACAAACAAATGGACTT 2160  
QY 2161 TATACTAGCCTCGGCCCATTTGGCACCCGTTACCTTACCCGTCCTCTGT 2209  
Db 2161 TATACTAGCCTCGGCCCATTTGGCACCCGTTACCTTACCCGTCCTCTGT 2209

RESULT 12  
AD226930  
ID AD226930 standard; DNA; 2214 BP.  
XX  
AC AD226930;  
XX  
DT 30-JUN-2005 (first entry)  
XX  
DE Adeno-associated virus DNA SEQ ID NO 80.  
XX  
KW rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;  
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;  
KW bacterial infection; cancer; ulcerative colitis; antineumatic;  
KW antihistatic; neuroprotective; antiinflammatory; antidiabetic;  
KW antipsoritic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;  
KW viricide; antibacterial; cytostatic; antiulcer; dermatological; ds.  
XX  
OS Adeno-associated virus.  
XX  
PN WO200503321-A2.  
XX  
PD 14-APR-2005.  
XX  
PP 30-SEP-2004; 2004WO-US028817.  
XX  
PR 30-SEP-2003; 2003US-0508226P.  
PR 29-APR-2004; 2004US-0566546P.  
XX  
PA (UTPE-) UNIV PENNSYLVANIA.  
XX  
XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;  
XX WPI; 2005-285437/29.  
XX  
XX New adeno-associated virus (AAV) clade comprising at least three AAV  
XX members, useful for preventing and/or treating arthritis, multiple  
XX sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial  
XX infection and cancer.  
XX  
XX Claim 50; SEQ ID NO 80; 569pp; English.  
XX  
XX The invention relates to an adeno-associated virus (AAV) clade comprising  
XX at least three AAV members, where each member of the AAV clade is  
XX phylogenetically related as determined using a Neighbor-Joining heuristic  
XX by a bootstrap value of at least 75 % per 1000 isolates and a Poisson  
XX correction distance measurement of no more than 0.05. The methods and  
XX compositions of the present invention are useful for the prevention  
XX and/or treatment of rheumatoid arthritis, multiple sclerosis,  
XX sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's  
XX disease, hemophilia, HIV, bacterial infection, cancer and ulcerative  
XX colitis. The present sequence represents an adeno-associated virus DNA.  
XX  
XX Sequence 2214 BP; 572 A; 658 C; 543 G; 441 T; 0 U; 0 Other;

Query Match 96.2%; Score 2124; DB 14; Length 2214;  
Best Local Similarity 97.8%; Pred. No. 0;  
Matches 2164; Conservative 0; Mismatches 45; Indels 3; Gaps 1;  
QY 1 ATGGCTGCGGATGGTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60  
Db 1 ATGGCTGCGGATGGTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60  
QY 61 GAGTGGTGGGATCTGAAACCTTGAGGCCCGGAAAGCCCAACACGACAAAGAGAGAC 120  
Db 61 GAGTGGTGGGATCTGAAACCTTGAGGCCCGGAAAGCCCAACACGACAAAGAGAGAC 120  
QY 121 GACGCGCGGGTCTGGTCTTCTGAGTACCTCGGACCTTCAACGGACTTCGAC 180  
Db 121 GACGCGCGGGTCTGGTCTTCTGAGTACCTCGGACCTTCAACGGACTTCGAC 180  
QY 181 AAGGGGAGCGCGTCAACGCGCGGACGACGCGGCTTCGAGACGACAAAGGCTTACGAC 240  
Db 181 AAGGGGAGCGCGTCAACGCGCGGACGACGCGGCTTCGAGACGACAAAGGCTTACGAC 240  
QY 241 CAGCAGCTCAAGCGGGTGACAATCCGTTACTGGGGTATAACCGACGCGGAGTTT 300  
Db 241 CAGCAGCTCAAGCGGGTGACAATCCGTTACTGGGGTATAACCGACGCGGAGTTT 300  
QY 301 CAGGAGCGTCTGCAAGAGATACGTTTGGGGGCACTTGGGGGAGAGTCTTCCAG 360  
Db 301 CAGGAGCGTCTGCAAGAGATACGTTTGGGGGCACTTGGGGGAGAGTCTTCCAG 360  
QY 361 GCCAAGAAAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTTAAAGACGGCTCCT 420  
Db 361 GCCAAGAAAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTTAAAGACGGCTCCT 420  
QY 421 GGAAGAAACGTCGGGTAGACGATCGCCACA---AGAGCCAGACTCTCTCTCGGGCATC 477  
Db 421 GGAAGAAAGACGCGGTAGAACCGTCACTCAGCGTTCTCCCGACTCTCTCAACGGGCATC 480  
QY 478 GCAAGACAGGCGCAGCGCCGCTTAAAGAGACTCAATTTTGGTTCAGACTGGGCACTCA 537  
Db 481 GCAAGAAAGCGCAGCGCCGCTTAAAGAGAGACTGAACTTTGGTTCAGACTGGGCACTCA 540  
QY 538 GAGTCAGTCCCGATCCACAACTCTCGGAGAACTCCAGCAACCCCGCTGCTGTGGGA 597  
Db 541 GAGTCAGTCCCGATCCACAACTCCAGCAACCCCGCTGCTGTGTGGGA 600  
QY 598 CTTACTCAATGGCTTCAGGCGGTGCGCAACAATGGCAGACAAATAAGAGGCGCCGAC 657  
Db 601 TCTGGTCAATGGCTTCAGGCGGTGCGCTCCAATGGCAGACAAATAAGAGGCGCCGAC 660  
QY 658 GAGTGGGTAAATGCTTCAGGAAATTTGGCATTTGCGATTCCCATGGCTGGGCGACAGTTC 717  
Db 661 GAGTGGGTAAATGCTTCAGGAAATTTGGCATTTGCGATTCCCATGGCTGGGCGACAGTTC 720  
QY 718 ATCCACACAGCAGCCGCACTTGGGCTTGGCCACTTACCAATAACCACTCTTACAAGCAA 777  
Db 721 ATCCACACAGCAGCCGCACTTGGGCTTGGCCACTTACCAATAACCACTCTTACAAGCAA 780  
QY 778 ATCTCCAGTGTCTCAACCGGGGCGCAGCAACCACTTCTCGGTTCAGGACCCGCC 837  
Db 781 ATCTCCAGTGTCTCAACCGGGGCGCAGCAACCACTTCTCGGTTCAGGACCCGCC 840  
QY 838 TGGGGGTATTTGATTTCAACAGATTCCATGCGCACTTTTCCACAGTGCAGTGGCAGCA 897  
Db 841 TGGGGGTATTTGATTTCAACAGATTCCATGCGCACTTTTCCACAGTGCAGTGGCAGCA 900  
QY 898 CTCATCAACAACTTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATC 957  
Db 901 CTCATCAACAACTTTGGGATTTCCGGGCCCAAGAGACTCAACTTCAAACTCTTCAACATC 960  
QY 958 CAAAGTCAAGAGGTCACGACGAATGATGGCGTCAACCACTCGCTTAATACCTTACCAGC 1017  
Db 961 CAAAGTCAAGAGGTCACGACGAATGATGGCGTCAACCACTCGCTTAATACCTTACCAGC 1020







121 GACGCGGGGCTGGTCTTCTGGCTACAGTACCTCGGACCCCTTCAACGGAATCGAC 180  
121 GACAGAGGGGCTTGTGCTTCTGGGTACAAGTACCTCGGACCCCTTCAACGGACTCGAC 180  
181 AAGGGGAGCCGCTCAACGCGCGGACGAGCGGCCCTCGAGCAGCAAGGCGCTACGAC 240  
181 AAGGAGAGCGGTCAAAGAGCAGACGCGCGGCCCTCGAGCAGCAAGGCTACGAC 240  
241 CAGCAGCTCAAGCGGGTGACAATTCGTACCTGCGGTATAACGCGCGAGCGGAGTTT 300  
241 CGGAGCTCGACAGCGGAGACAACCGGTACCTCAAGTACAACAACGCGCGAGCGGAGTTT 300  
301 CAGGAGCGTTCGCAAGAGATAGCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 360  
301 CAGGAGCGCTTAAGAAAGATAGCTTTTGGGGGCAACCTCGGGGAGCAGTCTTCCAG 360  
361 GCCAAGAAGCGGGTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT 420  
361 GCCAAGAAGCGGGTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTGAGACGGCTCCT 420  
421 GGAAGAAAGCTCCGCTGAGCAGTCCGCAAGAGCCAGACTCTCTCTCGGGCATCGGC 480  
421 GGAAGAAAGCTCCGCTGAGCAGTCCGCAAGAGCCAGACTCTCTCTCGGGCATCGGC 480  
481 AAGACAGCCAGAGCCCGCTAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 540  
481 AAGACAGCCAGAGCCCGCTAAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 540  
541 TCAGTCCCGGATCCACAACCTCTCGGAGAACCTTCAGCAACCCCGCTGCTGGGACCT 600  
541 TCAGTCCCGGATCCACAACCTCTCGGAGAACCTTCAGCAACCCCGCTGCTGGGACCT 600  
601 ACTAATAGCTTCAGCGGTGCGCAACCAATGGCAGACAATAAAGAGCGCGCAGCGA 660  
601 ACTAATAGCTTCAGCGGTGCGCAACCAATGGCAGACAATAAAGAGCGCGCAGCGA 660  
661 GTGGGTAAATGCTTCAGGAAATGGCAATGGCATTCGATTCACATGGCTGGGCGACAGTCA 720  
661 GTGGGTAAATGCTTCAGGAAATGGCAATGGCATTCGATTCACATGGCTGGGCGACAGTCA 720  
721 ACCACAGACACCGGACCTGGGCTTGGCCACCTACAATAACCACTCTACAAGCAAAATC 780  
721 ACCACAGACACCGGACCTGGGCTTGGCCACCTACAATAACCACTCTACAAGCAAAATC 780  
781 TCAGTGTCTTCAACGGGGGCGAGCAACCACTACTTTCGGCTACAGCAACCCCGCTGG 840  
781 TCAGTGTCTTCAACGGGGGCGAGCAACCACTACTTTCGGCTACAGCAACCCCGCTGG 840  
841 GGGTATTTGATTTCAACAGATTCACCTGCCACTTTTACCAGTGAAGTGGGAGGACTC 900  
841 GGGTATTTGATTTCAACAGATTCACCTGCCACTTTTACCAGTGAAGTGGGAGGACTC 900  
901 ATCAACAACAAATTTGGGATTCGGGCGGCAAGAGACTCAAATTTCAAACTCTTCAACATCCAA 960  
901 ATCAACAACAAATTTGGGATTCGGGCGGCAAGAGACTCAAATTTCAAACTCTTCAACATCCAA 960  
961 GTCAAGAGGTCAAGCAAGATATGGGTCAACCAATCGCTAATACCTTTACCAAGCAG 1020  
961 GTCAAGAGGTCAAGCAAGATATGGGTCAACCAATCGCTAATACCTTTACCAAGCAG 1020  
1021 GTTCAAGTCTTCTCGGACTCGGAGTACCAGTTCGGTACGTCTCGGCTCTGGGACCGAG 1080  
1021 GTTCAAGTCTTCTCGGACTCGGAGTACCAGTTCGGTACGTCTCGGCTCTGGGACCGAG 1080  
1081 GGTGCTCCTCCTCGGTTCCCGGGGAGCGTGTTCATGATTCGCAATACGGCTACCTGAGC 1140  
1081 GGTGCTCCTCCTCGGTTCCCGGGGAGCGTGTTCATGATTCGCAATACGGCTACCTGAGC 1140  
1141 CTCAACAATGGCAGCCAGCCGTTGGGACGTTTCATCTCTTTTACGCTGGGAATATTTCCCT 1200  
1141 CTCAACAATGGCAGCCAGCCGTTGGGACGTTTCATCTCTTTTACGCTGGGAATATTTCCCT 1200

RESULT 14

AD227053

ID AD227053 standard; DNA; 2211 BP.

XX

1201 TCTCAGATGCTGAGAAACCGGCAACAACTTTACCTTACGCTACACCTTTGAGGAAGTGCT 1260  
1201 TCTCAGATGCTGAGAAACCGGCAACAACTTTACCTTACGCTACACCTTTGAGGAAGTGCT 1260  
1261 TTCCAACAGAGCTACGCGCACAGCGAGCCCTGGACCGGCTGATGAATCTCTCTCATCGAC 1320  
1261 TTCCAACAGAGCTACGCGCACAGCGAGCCCTGGACCGGCTGATGAATCTCTCTCATCGAC 1320  
1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAAGTCCCAAAACAAGAC 1380  
1321 CAATACCTGTATTACCCGAACAGAACTCAAAATCAGTCCGGAAAGTCCCAAAACAAGAC 1380  
1381 TTGCTGTTTTAGCGTGGGCTCTCCAGCTGGCATGTCTTTCAGCCCAAAACCTGGTACT 1440  
1381 TTGCTGTTTTAGCGTGGGCTCTCCAGCTGGCATGTCTTTCAGCCCAAAACCTGGTACT 1440  
1441 GGACCTCTGTATTACCGGACAGCGGCTTTCTAAAACAAAAAACAGACAACCAACAGCAAT 1500  
1441 GGACCTCTGTATTACCGGACAGCGGCTTTCTAAAACAAAAAACAGACAACCAACAGCAAT 1500  
1501 TTTTACCTGGAATGGTGTCTTCAAAATATTAACCTCAATGGCGGTGAATCCATCATCAACCT 1560  
1501 TTTTACCTGGAATGGTGTCTTCAAAATATTAACCTCAATGGCGGTGAATCCATCATCAACCT 1560  
1561 GGCACCTGTATGGCTTCACAAAGACGACGAAGACAAGTTCTTTCCATGAGCGGTGTC 1620  
1561 GGCACCTGTATGGCTTCACAAAGACGACGAAGACAAGTTCTTTCCATGAGCGGTGTC 1620  
1621 ATGATTTTGGAAAAAGAGAGCGCGGAGCTTCAAACTGCAATTTGGCAAAATGTCAAT 1680  
1621 ATGATTTTGGAAAAAGAGAGCGCGGAGCTTCAAACTGCAATTTGGCAAAATGTCAAT 1680  
1681 ACAGAGAGAGAGAAATTTAAAGCCTAAACCTGTGGGCAACGAAAGATTTGGGACCGTG 1740  
1681 ACAGAGAGAGAGAAATTTAAAGCCTAAACCTGTGGGCAACGAAAGATTTGGGACCGTG 1740  
1741 CGAGTCAATTTCCAGCAGCAGCAGCAGACACCTCGACCGGAGATGTCATGCTATGGGA 1800  
1741 CGAGTCAATTTCCAGCAGCAGCAGCAGACACCTCGACCGGAGATGTCATGCTATGGGA 1800  
1801 GCATTTACCTGGCATGGTGTGGCAAGTAGAGACGTGTACCTGACGGGTCCCAATTTGGGCC 1860  
1801 GCATTTACCTGGCATGGTGTGGCAAGTAGAGACGTGTACCTGACGGGTCCCAATTTGGGCC 1860  
1861 AAAATTCCTCAACAGATGGACATTTTCAACCGGTCTCTCTTATGGGCGGCTTTGGACTC 1920  
1861 AAAATTCCTCAACAGATGGACATTTTCAACCGGTCTCTCTTATGGGCGGCTTTGGACTC 1920  
1921 AAGAACCCGCTCTCAGATCCTCATCAAAACACGCTGTTCCTGCGAATCCTCCGGCG 1980  
1921 AAGAACCCGCTCTCAGATCCTCATCAAAACACGCTGTTCCTGCGAATCCTCCGGCG 1980  
1981 GAGTTTTCAGCTCAAAAGTTTGTCTTCAATCATCACCCCAATATCTCCACAGGACAAGTGA 2040  
1981 GAGTTTTCAGCTCAAAAGTTTGTCTTCAATCATCACCCCAATATCTCCACAGGACAAGTGA 2040  
2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAG 2100  
2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAAATCCCGAAGTGCAG 2100  
2101 TACACATCCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGACAT 2160  
2101 TACACATCCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGACAT 2160  
2161 TATACTGAGCCTCGCCCCATTGGCACCCGTTACCTTTACCCGTCCTCTGT 2209  
2161 TATACTGAGCCTCGCCCCATTGGCACCCGTTACCTTTACCCGTCCTCTGT 2209

AC AD227053;  
XX 30-JUN-2005 (first entry)  
XX Adeno-associated virus DNA SEQ ID NO 203.  
DE  
XX rheumatoid arthritis; multiple sclerosis; sarcoidosis; diabetes;  
KW scleroderma; psoriasis; vasculitis; Crohn's disease; hemophilia; HIV;  
KW bacterial infection; cancer; ulcerative colitis; antirheumatic;  
KW antiarthritic; neuroprotective; antiinflammatory; antidiabetic;  
KW antipsoriatic; vasotropic; gastrointestinal-gen.; hemostatic; anti-HIV;  
XX virucide; antibacterial; cytostatic; antiulcer; dermatological; ds.  
OS Adeno-associated virus.  
XX WO2005033321-A2.  
XX  
XX 14-APR-2005.  
XX 30-SEP-2004; 2004WO-US028817.  
XX 30-SEP-2003; 2003US-0508226P.  
PR 29-APR-2004; 2004US-0566546P.  
XX (UYPE-) UNIV PENNSYLVANIA.  
PA  
XX Wilson JM, Gao G, Alvira MR, Vandenbergh LH;  
PI WPI; 2005-285437/29.  
XX  
XX New adeno-associated virus (AAV) clade comprising at least three AAV  
PT members, useful for preventing and/or treating arthritis, multiple  
PT sclerosis, diabetes, scleroderma, psoriasis, hemophilia, HIV, bacterial  
PT infection and cancer.  
XX  
XX Disclosure; SEQ ID NO 203; 569pp; English.  
PS  
XX The invention relates to an adeno-associated virus (AAV) clade comprising  
CC at least three AAV members, where each member of the AAV clade is  
CC phylogenetically related as determined using a Neighbor-Joining heuristic  
CC by a bootstrap value of at least 75 % per 1000 isolates and a Poisson  
CC correction distance measurement of no more than 0.05. The methods and  
CC compositions of the present invention are useful for the prevention  
CC and/or treatment of rheumatoid arthritis, multiple sclerosis,  
CC sarcoidosis, diabetes, scleroderma, psoriasis, vasculitis, Crohn's  
CC disease, hemophilia, HIV, bacterial infection, cancer and ulcerative  
CC colitis. The present sequence represents an adeno-associated virus DNA.  
XX  
XX Sequence 2211 BP; 576 A; 647 C; 547 G; 441 T; 0 U; 0 Other;  
SQ  
Query Match 95.3%; Score 2105; DB 14; Length 2211;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 2144; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
QY 1 ATGGCTGCCGATGGTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60  
DB 1 ATGGCTGCCGATGGTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC 60  
QY 61 GAGTGTGGGACTTGAACCTCGAGCCCCGAAAGCCCAAGCCAAACAGCAAGAGCAGGAC 120  
DB 61 GAGTGTGGGACTTGAACCTCGAGCCCCGAAAGCCCAAGCCAAACAGCAAGAGCAGGAC 120  
QY 121 GACGGCCGGGCTGGTGTCTTCTGGCTACAAGTACCTCGGACCCCTTCACGACCTCGAC 180  
DB 121 GACGGCCGGGCTGGTGTCTTCTGGCTACAAGTACCTCGGACCCCTTCACGACCTCGAC 180  
QY 181 AAGGGGAGCCCGTCAACGCGGCGAGCAGCGGCCCTCGAGACCAAGAGCCCTACGAC 240  
DB 181 AAGGGGAGCCCGTCAACGCGGCGAGTGCAGCGGCCCTCGAGACCAAGAGCCCTACGAC 240  
QY 241 CAGCAGCTCAAGCGGGTGACAAATCCGTACTTCGGGTATACCAACGCGGCGGAGTTT 300  
DB 241 CAGCAGCTCAAGCGGGTGACAAATCCGTACTTCGGGTATACCAACGCGGCGGAGTTT 300

QY 301 CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGCAACCTCGGCGAGCAGTCTTCCAG 360  
DB 301 CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGCAACCTCGGCGAGCAGTCTTCCAG 360  
QY 361 GCCAAGAGCGGTTCTCGAACCTCTCGGTCTGGTTGAGGAGCGCTTAAGACGGCTCT 420  
DB 361 GCCAAGAGAGGGTTCTCGAACCTTTTGGTCTGGTTGAGGAGGGTCTTAAGACGGCTCT 420  
QY 421 GGAAAGAAACGTCGGTAGAGCAGTCGCCAAGAGCCAGACTCTCTCTCGGCGCATCGGC 480  
DB 421 GGAAAGAAACGTCGGTAGAGCAGTCGCCAAGAGCCAGACTCTCTCTCGGCGCATTGGC 480  
QY 481 AAGACAGCCAGCAGCCCGCTTAAAGAGAGACTCAATTTTGGTTCAGCTGGCGACTCAGAG 540  
DB 481 AAGACAGCCAGCAGCCCGCTTAAAGAGAGACTCAATTTTGGTTCAGACTGGCGACTCAGAG 540  
QY 541 TCAGTCCCGATCCACNAACCTCTCGGAGAACCTCCAGCAACCCCGCTCTGTGGGACCT 600  
DB 541 TCAGTCCCGACCCCAACCTCTCGGAGAACCTCCAGCAACCCCGCTCTGTGTGGGACCT 600  
QY 601 ACTACAAATGGCTTCAGGGCGGTGGCGCACCAATGGCAGACAATAAGAGCGCCGACGGA 660  
DB 601 ACTACAAATGGCTTCAGGGCGGTGGCGCACCAATGGCAGACAATAAGAGCGCCGACGGA 660  
QY 661 GTGGGTAAATGCTCAGGAAATTTGGCAATTCGCAATTCACATGGTGGCGCAGAGTCAATC 720  
DB 661 GTGGGTAAATGCTCAGGAAATTTGGCAATTTGGCAATTCGCAATTCACATGGTGGCGCAGAGTCAATC 720  
QY 721 ACCACAGCAGCCGACCTGGGCTTGGCCACCTACNAATAACCACTCTACAGCAAAATC 780  
DB 721 ACCACAGCAGCCGACCTGGGCTTGGCCACCTTAAACACCACTCTACAGCAAAATC 780  
QY 781 TCCAGTGTCTCAAGCGGGCCAGCAACCACTACTCTCGGCTACAGCAGCCCGCTGG 840  
DB 781 TCCAGTGTCTCAAGCGGGCCAGCAACCACTACTCTCGGCTACAGCAGCCCGCTGG 840  
QY 841 GGGTATTTGATTTCAACAGATTCCAATGCGCACTTTTCAACAGTGTGGCAGCGACTC 900  
DB 841 GGGTATTTGATTTCAACAGATTCCAATGCGCACTTTTCAACAGTGTGGCAGCGACTC 900  
QY 901 ATCAACAACAAATGGGGATTTCGGCCCAAGAGACTCAACTTCAACTCTTCAACATCCAA 960  
DB 901 ATCAACAACAAATGGGGATTTCGGCCCAAGAGACTCAACTTCAAGTCTTCAACATCCAA 960  
QY 961 GTCAAGGAGGTCAACGAGCAATGATGGCGTCAACCACTCGCTAAATACCTTACAGCAGC 1020  
DB 961 GTCAAGGAGGTCAACGAGCAATGATGGCGTCAACCACTCGCTAAATACCTTACAGCAGC 1020  
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCAAGTTCCTGATCGTCTCGGCTCTGGCACCAG 1080  
DB 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCAAGTTCCTGATCGTCTCGGCTCTGGCACCAG 1080  
QY 1081 GGTGCTCTCTCGGTTCCGGCGGAGCTGTTCAATGATTCGCGCAATACGGCTACCTGACG 1140  
DB 1081 GGTGCTCTCTCGGTTCCGGCGGAGCTGTTCAATGATTCGCGAGTACGGCTACCTGACG 1140  
QY 1141 CTCACAATGGGAGCAAGCCGTTGGGAGCTTCACTCTTTTACTGCTCGGAATATTTCCCT 1200  
DB 1141 CTCACAATGGGAGCAAGCCGTTGGGAGCTTCACTCTTTTACTGCTCGGAATATTTCCCA 1200  
QY 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTACACCTTTGAGGAGTGCCT 1260  
DB 1201 TCGCAGATGCTGAGAACGGGCAATACTTTTACCTTCAGCTACACCTTCGAGGAGTGCCT 1260  
QY 1261 TTTCCACAGAGCTACCGGCAAGCAGGCTGGACCGGCTGATGATCTCTCTCATCGAC 1320  
DB 1261 TTTCCACAGAGCTACCGGCAAGCAGGCTGGACCGGCTGATGATCTCTCTCATCGAC 1320  
QY 1321 CAATACCTGTATTACCTGAGCAAGAACTCAAAATCAGTCCGAGAGTCCCAAAACAGGAC 1380  
DB 1321 CAGTACTGTATTACCTGAGCAAGAACTCAAAATCAGTCCGAGAGTCCCAAAACAGGAC 1380

QY 1381 TTGCTGTTTACGGCTGGGTTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAAACCTGGCTACCT 1440  
DB |||||  
QY 1381 TTGCTGTTTACGGCTGGGTTCTCCAGCTGGCATGTCTGTTTCAGCCCAAAAACCTGGCTACCT 1440  
DB |||||  
QY 1441 GGAACCTGTTTATCGGCAGCAGCGCGTTTCTTAAAAAAGAGACAAACAAACAGCAAT 1500  
DB |||||  
QY 1441 GGAACCTGTTTATCGGCAGCAGCGCGTTTCTTAAAAAAGAGACAAACAAACAGCAAT 1500  
DB |||||  
QY 1501 TTATCTGGAAGTGGTGTCTTAAAAATATAAATCTCAATGGCGGTGAATTCATCATCAACCT 1560  
DB |||||  
QY 1501 TTATCTGGAAGTGGTGTCTTAAAAATATAAATCTCAATGGCGGTGAATTCATCATCAACCT 1560  
DB |||||  
QY 1561 GGCACCTGCTATGGCTCTCACAAAGACGACGAGCAAGTCTTCTTCCCATGAGCGGTGTC 1620  
DB |||||  
QY 1561 GGCACCTGCTATGGCTCTCACAAAGACGACGAGCAAGTCTTCTTCCCATGAGCGGTGTC 1620  
DB |||||  
QY 1621 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAACACATGCTATGGCAATGTGCATGATT 1680  
DB |||||  
QY 1621 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAACACATGCTATGGCAATGTGCATGATT 1680  
DB |||||  
QY 1681 ACAGACGAAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTGGGACCGTG 1740  
DB |||||  
QY 1681 ACAGACGAAGAGAAATTAAGCCACTAAACCTGTGGCCACCGAAAGATTGGGACCGTG 1740  
DB |||||  
QY 1741 CGAGTCAATTTCCAGAGCAGCAGCAGACACCTTGGACCGGAGATGTGCATGCTATGGA 1800  
DB |||||  
QY 1741 CGAGTCAATTTCCAGAGCAGCAGCAGACACCTTGGACCGGAGATGTGCATGCTATGGA 1800  
DB |||||  
QY 1801 GCATTAACCTGGCATGTGGCAAGATAGAGAGCTGACTGCAGGTTCCCATTTGGGCG 1860  
DB |||||  
QY 1801 GCATTAACCTGGCATGTGGCAAGATAGAGAGCTGACTGCAGGTTCCCATTTGGGCG 1860  
DB |||||  
QY 1861 AAAATTCCTCACACAGATGGACACTTTCACCCGCTCTCTCTTATGGGCGGCTTTGGACTC 1920  
DB |||||  
QY 1861 AAAATTCCTCACACAGATGGACACTTTCACCCGCTCTCTCTTATGGGCGGCTTTGGACTC 1920  
DB |||||  
QY 1921 AAGAACCGGCTCTCAGATCTCATCAAAACACGCTGTTCCTCTGCGAATCTCTCGGCG 1980  
DB |||||  
QY 1921 AAGAACCGGCTCTCAGATCTCATCAAAACACGCTGTTCCTCTGCGAATCTCTCGGCG 1980  
DB |||||  
QY 1981 GAGTTTTCAGCTACAAAGTTTGTCTCATTCATCACCCTCAATATCTCCACAGGACAAGTGAGT 2040  
DB |||||  
QY 1981 GAGTTTTCAGCTACAAAGTTTGTCTCATTCATCACCCTCAATATCTCCACAGGACAAGTGAGT 2040  
DB |||||  
QY 2041 GTGAAAATTAATGGAGCTGCAAGAAAGAAACAGCAAGCGCTGGAAATCCGGAAGTGCAG 2100  
DB |||||  
QY 2041 GTGGAATTAATGGAGCTGCAAGAAAGAAACAGCAAGCGCTGGAAATCCGGAAGTGCAG 2100  
DB |||||  
QY 2101 TACACATCCAAATATGCAAAATCTGCCAAGTTGATTTTACTGTGGACAAATGGACTT 2160  
DB |||||  
QY 2101 TATACATCTAACTATGCAAAATCTGCCAAGTTGATTTTACTGTGGACAAATGGACTT 2160  
DB |||||  
QY 2161 TATACTGAGCTCGCCCGCATTTGGACCCGTTACCTTACCGTCCCGCTGT 2209  
DB |||||  
QY 2161 TATACTGAGCTCGCCCGCATTTGGACCCGTTACCTTACCGTCCCGCTGT 2209  
DB |||||

## RESULT 15

ADW39402

ID ADM39402 standard; DNA; 4239 BP.

XX AC

XX DT

XX XX

XX 24-MAR-2005 (first entry)

XX Adeno-associated virus 2 (AAV2) rep-AAV1 cap fusion DNA - SEQ ID 5.

XX genetically engineered microorganism; vector; rep; cap; ds.

XX Adeno-associated virus 2.

XX Chimeric.

XX Unidentified.

PN CN1461805-A.  
XX PD 17-DEC-2003.  
XX PF 27-MAY-2002; 2002CN-00117965.  
XX PR 27-MAY-2002; 2002CN-00117965.  
XX PA (BENY-) BENYUAN ZHENGYANG GENETIC TECHNOLOGY CO.  
XX PI Wu X, Cao H, Dong X;  
XX WPI; 2004-248658/24.  
XX Structure of a recombinant herpes simplex virus and its use.  
XX Claim 5; SEQ ID NO 5; 102pp; Chinese.  
XX The invention relates to a novel recombinant herpes simplex virus (rHSV) whose genome is respectively inserted by the Adeno-associated virus (AAV) rep gene and cap gene. The AAV genes may be derived from viruses AAV1, AAV2, AAV3, AAV4, AAV5 and AAV6. The methods of the invention may be useful for generating recombinant vectors with high efficiency and no need of reconfiguration. The current sequence is that of the Adeno-associated virus 2 (AAV2) rep-Adeno-associated virus 1 (AAV1) cap fusion DNA (SEQ ID 5) of the invention.  
XX Sequence 4239 BP; 1084 A; 1156 C; 1130 G; 869 T; 0 U; 0 Other;  
QY Query Match 95.3%; Score 2105; DB 13; Length 4239;  
DB Best Local Similarity 97.1%; Pred. No. 0;  
Matches 2144; Conservative 0; Mismatches 65; Indels 0; Gaps 0;  
QY 1 ATGGCTGCGGATGGTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCAATTGCG 60  
DB 2021 ATGGCTGCGGATGGTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGCAATTGCG 2080  
QY 61 GAGTGTGGGACTTGAACCTGGAGCCCGGAGCCCAAGCCCAAGCCCAAGCAAGAGGAGAC 120  
DB 2081 GAGTGTGGGACTTGAACCTGGAGCCCGGAGCCCAAGCCCAAGCCCAAGCAAGAGGAGAC 2140  
QY 121 GAGGCGCGGGTCTGGTCTTCTGGCTACAGTACCTCGGACCTTCAACGGACTCGAC 180  
DB 2141 GAGGCGCGGGTCTGGTCTTCTGGCTACAGTACCTCGGACCTTCAACGGACTCGAC 2200  
QY 181 AAGGGGAGCGCGTCAACCGCGGACGCGAGCGGCTCGAGCAGCAAGGCTTACGAC 240  
DB 2201 AAGGGGAGCGCGTCAACCGCGGAGTGAAGCGGCTCGAGCAGCAAGGCTTACGAC 2260  
QY 241 CAGCAGCTCAAGCGGGTGACAATCCGTACTCGGTATAACACCGCGAGCCGAGTTT 300  
DB 2261 CAGCAGCTCAAGCGGGTGACAATCCGTACTCGGTATAACACCGCGAGCCGAGTTT 2320  
QY 301 CAGGAGCGTCTGCAAGAGATACGTTTGGGGGCAACCTCGGGGAGGAGTCTTCCAG 360  
DB 2321 CAGGAGCGTCTGCAAGAGATACGTTTGGGGGCAACCTCGGGGAGGAGTCTTCCAG 2380  
QY 361 GCCAAGAAAGCGGTTCTCGAACTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCT 420  
DB 2381 GCCAAGAAAGCGGTTCTCGAACTCTGGTCTGGTTGAGGAAGGCTAAGACGGCTCT 2440  
QY 421 GGAAGAAAGCGTCCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCTCGGGCATTCGC 480  
DB 2441 GGAAGAAAGCGTCCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCTCGGGCATTCGC 2500  
QY 481 AAGACAGGCCAGCGCCGCTTAAAGAGACTCAATTTTGGTCAGACTGGCCACTCAGAG 540  
DB 2501 AAGACAGGCCAGCGCCGCTTAAAGAGACTCAATTTTGGTCAGACTGGCCACTCAGAG 2560  
QY 541 TCAGTCCCCGATCCCAAACTCTCGGAGAACTTCCAGCAACCCCGCTGCTGTGGACCT 600  
DB 2561 TCAGTCCCCGATCCCAAACTCTCGGAGAACTTCCAGCAACCCCGCTGCTGTGGACCT 2620

QY 601 ACTACAAATGGCTTTCAGCGGTTGGCGCACCAATGGCAGACAATAACGAAGCGCCGACGGA 660  
DB 2621 ACTACAAATGGCTTTCAGCGGTTGGCGCACCAATGGCAGACAATAACGAAGCGCCGACGGA 2680  
QY 661 GTGGGTAAATGGCTCAGGAAATTTGGCAATTTGGATTTCCACATGGCTGGCGGCGACAGATCATC 720  
DB 2681 GTGGGTAAATGGCTCAGGAAATTTGGCAATTTGGATTTCCACATGGCTGGCGGCGACAGATCATC 2740  
QY 721 ACCACAGACCCCGACCTTGGGCTTTGGCCACCTACAAATAACCACTCTTACAAAGCAAAATC 780  
DB 2741 ACCACAGACCCCGACCTTGGGCTTTGGCCACCTTACAAATAACCACTCTTACAAAGCAAAATC 2800  
QY 781 TCCAGTGTCTTCAACGGGGCCGACGAAACGACAACTACTTTCGGCTTACAGCAACCCCTGG 840  
DB 2801 TCCAGTGTCTTCAACGGGGCCGACGAAACGACAACTACTTTCGGCTTACAGCAACCCCTGG 2860  
QY 841 GGGTATTTTGAATTTCAACAGATTTCCACTGGCACTTTTACACAGCTGACTGGCAGGCACTC 900  
DB 2861 GGGTATTTTGAATTTCAACAGATTTCCACTGGCACTTTTACACAGCTGACTGGCAGGCACTC 2920  
QY 901 ATCAACAAATTTGGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA 960  
DB 2921 ATCAACAAATTTGGGGATTTCCGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA 2980  
QY 961 GTCAAGGAGGTCAACGAGATGATGGCGTCAACCACTCGCTAATAACCTTACCAAGCAG 1020  
DB 2981 GTCAAGGAGGTCAACGAGATGATGGCGTCAACCACTCGCTAATAACCTTACCAAGCAG 3040  
QY 1021 GTTCAAGTCTTTCGGACTCGGAGTACAGCTTCCGTAAGTCTCGGCTTGGCGACCAAG 1080  
DB 3041 GTTCAAGTCTTTCGGACTCGGAGTACAGCTTCCGTAAGTCTCGGCTTGGCGACCAAG 3100  
QY 1081 GGTGCTCTCTCCGTTCCGGCGGAGCTGTCATGATTTCCGCAATACGGCTACCTGACG 1140  
DB 3101 GGTGCTCTCTCTCCGTTCCGGCGGAGCTGTCATGATTTCCGCAATACGGCTACCTGACG 3160  
QY 1141 CTCAACAAATGGCAGCAAGCGTGGGACGTTTCACTCTTTTACTGCTGGGAATTTTCCCT 1200  
DB 3161 CTCAACAAATGGCAGCAAGCGTGGGACGTTTCACTCTTTTACTGCTGGGAATTTTCCCA 3220  
QY 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTTACCTTTTGGAGGAGTGCCT 1260  
DB 3221 TCGCAGATGCTGAGAACGGGCAACAACTTTTACCTTCAGCTTACCTTTTGGAGGAGTGCCT 3280  
QY 1261 TTCCACAGCAGCTACCGGCGACAGCCGAGCCTGGACGGGTGATGATCTCTCATCGAC 1320  
DB 3281 TTCCACAGCAGCTACCGGCGACAGCCGAGCCTGGACGGGTGATGATCTCTCATCGAC 3340  
QY 1321 CAATACCTGTATTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAAAGGAC 1380  
DB 3341 CAGTACCTGTATTACCTGAAACAGAACTCAGNATCAGTCCGGAAGTGCCTCAAAACAAAGGAC 3400  
QY 1381 TTGCTGTGTTAGCGGTGGCTTCCAGCTGGCATGCTGTTTTCAGCCCAAAACTGGGTACCT 1440  
DB 3401 TTGCTGTGTTAGCGGTGGCTTCCAGCTGGCATGCTGTTTTCAGCCCAAAACTGGGTACCT 3460  
QY 1441 GGAACCTGTGTTAGCGGAGCGGCTTTTAAACCAAAACAGACACACACACACACACAC 1500  
DB 3461 GGAACCTGTGTTAGCGGAGCGGCTTTTAAACCAAAACAGACACACACACACACACAC 3520  
QY 1501 TTTTACCTGGAAGTGGTCTTCAAAATAAATCACTCAATGGGGGTGAATCCATCAACCCCT 1560  
DB 3521 TTTTACCTGGAAGTGGTCTTCAAAATAAATCACTTAAATGGGGGTGAATCTATATCAACCCCT 3580  
QY 1561 GGCACGTGCTATGGCTTACAAAGAGCGAGCAAGTCTTTTCCCATGAGCGGTGTC 1620  
DB 3581 GGCACGTGCTATGGCTTACAAAGAGCGAGCAAGTCTTTTCCCATGAGCGGTGTC 3640  
QY 1621 ATGATTTTGGAAAAAGAGCGGCGGAGCTTCAAACTGATGATTTGGCAATGTCATGAT 1680  
DB 3641 ATGATTTTGGAAAAAGAGCGGCGGAGCTTCAAACTGATGATTTGGCAATGTCATGAT 3700  
QY 1681 ACAGACGAGAGGAAATTAAGCCACTAACCTCTGGCCACCGGAAAGATTTGGGACCGTG 1740

DB 3701 ACAGACGAGAGGAAATCAAAGCCACTAACCCCGTGGCCACCAGAAAGATTTGGGACTGTG 3760  
QY 1741 GCAGTCAATTTCCAGAGCAGCAGCAGACCCCTGGCCACCGAGATGTGCAATGCTATGGGA 1800  
DB 3761 GCAGTCAATTTCCAGAGCAGCAGCAGACCCCTGGCCACCGAGATGTGCAATGCTATGGGA 3820  
QY 1801 GCATTAACCTGGCATGGTGGCAAGATAGAGACGTGTACTCGAGGGTCCCATTTGGGCC 1860  
DB 3821 GCCTTTACCTGGAATGGTGGCAAGACAGAGACGTATACCTGCAGGGTCTTATTTGGGCC 3880  
QY 1861 AAAATTTCTCACACAGATGGACACTTTTACCCGCTCTCTTATATGGGCGGCTTTGACTC 1920  
DB 3881 AAAATTTCTCACACAGATGGACACTTTTACCCGCTCTCTTATATGGGCGGCTTTGACTC 3940  
QY 1921 AAGAAACCCGCTCTCACAGATCCTCATCAAAACACAGCCCTGTTCTTCCGAATCCTCCGGCG 1980  
DB 3941 AAGCAACCCGCTCTCACAGATCCTCATCAAAACACAGCCCTGTTCTTCCGAATCCTCCGGCA 4000  
QY 1981 GAGTTTTCAGCTACAAAGTTTCTTTCATCATCAACCAATACCTTCCACAGACAAGTGAGT 2040  
DB 4001 GAGTTTTCAGCTACAAAGTTTCTTTCATCATCAACCAATACCTTCCACAGACAAGTGAGC 4060  
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGGAATCCCGAAGTGCAG 2100  
DB 4061 GTGGAGATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGGAATCCCGAAGTGCAG 4120  
QY 2101 TACACATCCAATTTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGAATT 2160  
DB 4121 TATACATCTAATCTATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAAATGGAATT 4180  
QY 2161 TATACTGAGCCTCGCCCCCATTTGGCACCCGTTACCTTACCCGTCCTCTGT 2209  
DB 4181 TATACTGAGCCTCGCCCCCATTTGGCACCCGTTACCTTACCCGTCCTCTGT 4229

Search completed: November 28, 2005, 19:55:43

Job time : 1039.63 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 12:51:27 ; Search time 8072.36 seconds  
(without alignments)  
15555.198 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_2223\_4431  
Perfect score: 2209  
Sequence: 1 atggctgcgatgttatct.....ttacctaccgtccctgt 2209

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 588141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.in.\*  
3: gb.env.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pr.\*  
9: gb.roi.\*  
10: gb.sts.\*  
11: gb.sy.\*  
12: gb.un.\*  
13: gb.vi.\*  
14: gb.htg.\*  
15: gb.pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2209	100.0	2211	6	BD242771 Adeno-ass
2	2209	100.0	2211	6	CS073491 Sequence
3	2209	100.0	2211	6	CS073614 Sequence
4	2209	100.0	2211	6	AR562503 Sequence
5	2209	100.0	4718	6	BD242766 Adeno-ass
6	2209	100.0	4718	6	AR562498 Sequence
7	2209	100.0	4718	6	AX753251 Sequence
8	2209	100.0	4718	13	AF063497 Adeno-ass
9	2202.6	99.7	7447	6	AR527492 Sequence
10	2196.2	99.4	2211	6	CS073490 Sequence
11	2196.2	99.4	2211	13	AY530611 Adeno-ass
12	2124	96.2	2214	6	CS073492 Sequence
13	2124	96.2	2214	13	AY530606 Adeno-ass
14	2109.8	95.5	2211	6	CS073493 Sequence
15	2109.8	95.5	2211	13	AY530607 Adeno-ass
16	2105	95.3	2211	6	CS073615 Sequence
17	2105	95.3	4683	13	AF028704 Adeno-ass
18	2101.8	95.1	4683	6	BD242775 Adeno-ass

19	2101.8	95.1	4683	6	AR562507 Sequence
20	2100.2	95.1	2211	6	CS073494 Sequence
21	2100.2	95.1	2211	13	AY530609 Adeno-ass
22	1798	81.4	1800	6	BD242772 Adeno-ass
23	1798	81.4	1800	6	AR562504 Sequence
24	1639	74.2	2214	6	CS073452 Sequence
25	1639	74.2	2214	13	AY530567 Adeno-ass
26	1620	73.3	2214	6	CS073453 Sequence
27	1616.8	73.2	2211	13	AY242997 Non-human
28	1616.8	73.2	3122	6	AX753290 Sequence
29	1616.6	73.2	2214	6	CS073449 Sequence
30	1616.6	73.2	2214	13	AY530568 Adeno-ass
31	1615.2	73.1	3117	6	AX753288 Sequence
32	1615.2	73.1	3122	6	AX753287 Sequence
33	1613.6	73.0	3121	6	AX753289 Sequence
34	1611.8	73.0	2214	6	CS073625 Sequence
35	1611.8	73.0	4721	6	CQ972062 Sequence
36	1611.8	73.0	4721	6	CS073592 Sequence
37	1611.8	73.0	4721	6	AX753246 Sequence
38	1611.8	73.0	4721	13	AF513851 Adeno-ass
39	1607	72.7	2214	6	CS073447 Sequence
40	1607	72.7	2214	6	CS073431 Sequence
41	1605.4	72.7	2214	6	CS073448 Sequence
42	1605.4	72.7	2214	13	AY243020 Non-human
43	1605.4	72.7	2217	6	CS073454 Sequence
44	1605.4	72.7	3127	6	AX753260 Sequence
45	1603.8	72.6	2214	6	CS073444 Sequence

ALIGNMENTS

RESULT 1

BD242771

LOCUS

DEFINITION

Adeno-associated virus serum type 1 nucleic acid sequence, vector

and host cell containing the same.

2211 bp DNA linear PAT 17-JUL-2003

ACCESSION

BD242771.1

VERSION

JP 2002529098-A/6.

KEYWORDS

unidentified

SOURCE

unclassified

ORGANISM

unclassified.

REFERENCE

1 (bases 1 to 2211)

AUTHORS

Wilson,J.M. and Xiao,W.

TITLE

Adeno-associated virus serum type 1 nucleic acid sequence, vector

and host cell containing the same

Patent: JP 2002529098-A 6 10-SEP-2002;

JOURNAL

THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA

COMMENT

OS AAV-1

PN JP 2002529098-A/6

PD 10-SEP-2002

PF 02-NOV-1999 JP 2000581227

PR 05-NOV-1998 US 60/107114

PI JAMES M WILSON, WEIDONG XIAO

PC C12N15/09,A61K31/711,A61P43/00,C12N1/15,C12N1/19, PC

C12N1/21,

PC C12N5/10,C12P21/02//A61K35/76,C12N15/00,C12N5/00 CC

Adeno-associated virus serum type 1 nucleic acid sequence, CC

vector and host

CC cell containing the same

FT Key

Location/Qualifiers

(1)..(2208).

FEATURES

source

/organism="unidentified"

/mol\_type="genomic DNA"

/db\_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 2209; DB 6; Length 2211;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Qy	1	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC	60
Db	1	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC	60
Qy	61	GAGTGTGGGACTTTGAACCTCTGGAGCGCCCGAAGCCCAAAGCCAAACGCAAAAGACGAGAC	120
Db	61	GAGTGTGGGACTTTGAACCTCTGGAGCGCCCGAAGCCCAAAGCCAAACGCAAAAGACGAGAC	120
Qy	121	GA CGGC CGGGTCTGGTGCTTCTCTGGGTACAAGTACTCGAGACCTTCAACGCACTCGAC	180
Db	121	GA CGGC CGGGTCTGGTGCTTCTCTGGGTACAAGTACTCGAGACCTTCAACGCACTCGAC	180
Qy	181	AAGGGGAGCCCGTTCAACGCGGGGAGCGCAGCGGCCCTCGAGCAGCAAAAGGCTACGAC	240
Db	181	AAGGGGAGCCCGTTCAACGCGGGGAGCGCAGCGGCCCTCGAGCAGCAAAAGGCTACGAC	240
Qy	241	CAGCAGCTCAAGCGGGGTGACATCCGTACTCTGGGTATAACCAACGCGAGCGGAGTTT	300
Db	241	CAGCAGCTCAAGCGGGGTGACATCCGTACTCTGGGTATAACCAACGCGAGCGGAGTTT	300
Qy	301	CAGGAGCGTCTGCAAGGAAGTACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
Db	301	CAGGAGCGTCTGCAAGGAAGTACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
Qy	361	GCCAAAGCGGGTCTCGAAACCTCTCGGTCTGGTTGAGAAAGCGCTAAGACGGCTCCT	420
Db	361	GCCAAAGCGGGTCTCGAAACCTCTCGGTCTGGTTGAGAAAGCGCTAAGACGGCTCCT	420
Qy	421	GGAAGAAACGTCCGGTAGAGCAGTCCGCCACAGAGCCAGACCTCTCTCGGGCATCGGC	480
Db	421	GGAAGAAACGTCCGGTAGAGCAGTCCGCCACAGAGCCAGACCTCTCTCGGGCATCGGC	480
Qy	481	AAGACAGCGCAGCAGCCCGCTTAAAGAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG	540
Db	481	AAGACAGCGCAGCAGCCCGCTTAAAGAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG	540
Qy	541	TCAGTCCCGATCCACAACCTCTCGGAGAACCTTCAGCAACCCCGCTGTGTGGGACCT	600
Db	541	TCAGTCCCGATCCACAACCTCTCGGAGAACCTTCAGCAACCCCGCTGTGTGGGACCT	600
Qy	601	ACTACAATGGCTTCAGCGGGTGGCGCACCAATGGCAGACAAATACGAAGCGCCGACGGA	660
Db	601	ACTACAATGGCTTCAGCGGGTGGCGCACCAATGGCAGACAAATACGAAGCGCCGACGGA	660
Qy	661	GTGGGTAAATGCCTCAGGAAATTTGGCAATTCGATTCACATGGCTGGCGCAGAGTCATC	720
Db	661	GTGGGTAAATGCCTCAGGAAATTTGGCAATTCGATTCACATGGCTGGCGCAGAGTCATC	720
Qy	721	ACCACGACACCGGACCTGGGCTTGGCCACCTTACAATACCACTCTACAGCAATTC	780
Db	721	ACCACGACACCGGACCTGGGCTTGGCCACCTTACAATACCACTCTACAGCAATTC	780
Qy	781	TCCAGTGTCTCAACGGGGGCGAGCAACGAACCACTACTTTCGGCTACAGCACCCCTGG	840
Db	781	TCCAGTGTCTCAACGGGGGCGAGCAACGAACCACTACTTTCGGCTACAGCACCCCTGG	840
Qy	841	GGGTATTTTGATTTCAACAGATTTCCACTGGCCATTTTTCACACGTGACTGGCAGCACTC	900
Db	841	GGGTATTTTGATTTCAACAGATTTCCACTGGCCATTTTTCACACGTGACTGGCAGCACTC	900
Qy	901	ATCAACAACAAATTTGGGATTTCCGGCCCAAGNAGACTCAACTTCAAACTTTCACATCCAA	960
Db	901	ATCAACAACAAATTTGGGATTTCCGGCCCAAGNAGACTCAAACTTTCACATCCAA	960
Qy	961	GTCAAGGAGGTTCAGCGAATGATGGGGTCAACAACCTCGCTAATAAACCCTTACGACAG	1020
Db	961	GTCAAGGAGGTTCAGCGAATGATGGGGTCAACAACCTCGCTAATAAACCCTTACGACAG	1020
Qy	1021	GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCTCGTACGTCTCGGCTCTGGGCAACAG	1080
Db	1021	GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCCTCGTACGTCTCGGCTCTGGGCAACAG	1080

Qy	1081	GGCTGCTCCCTCCGTTCCGGCGGACGTTTCATGATTCCTCGCAATACCGCTACCTGACG	1141
Db	1081	GGCTGCTCCCTCCGTTCCGGCGGACGTTTCATGATTCCTCGCAATACCGCTACCTGACG	1140
Qy	1141	CTCAACAATGGCAGCCAAAGCCGTGGACGTTTCATCCTTTTACTGCTCTGGAAATATTTCCCT	1200
Db	1141	CTCAACAATGGCAGCCAAAGCCGTGGACGTTTCATCCTTTTACTGCTCTGGAAATATTTCCCT	1200
Qy	1201	TCTCAGATGCTGAGAACGGGCAACAATTTACCTTCAGCTTACACCTTTTGAGGAAGTGCCT	1260
Db	1201	TCTCAGATGCTGAGAACGGGCAACAATTTACCTTCAGCTTACACCTTTTGAGGAAGTGCCT	1260
Qy	1261	TTCCACAGCAGCTACGCGCACAGCCAGAGCCTCGACCGCGCTGATGAATCCTCTCATCGAC	1320
Db	1261	TTCCACAGCAGCTACGCGCACAGCCAGAGCCTCGACCGCGCTGATGAATCCTCTCATCGAC	1320
Qy	1321	CAATACCTGTATTTACCTGAACAGAACTCAAAAATCAGTCCGGAAAGTGCCTCAAAACAAGGAC	1380
Db	1321	CAATACCTGTATTTACCTGAACAGAACTCAAAAATCAGTCCGGAAAGTGCCTCAAAACAAGGAC	1380
Qy	1381	TTGCTGTTTAGCCGTGGCTCTCAGCTGGCATGTCTGTCTCAGCGCCAAAACTGGCTACCT	1440
Db	1381	TTGCTGTTTAGCCGTGGCTCTCAGCTGGCATGTCTGTCTCAGCGCCAAAACTGGCTACCT	1440
Qy	1441	GGACCTGTATTCCGCAGCAGCGCTTTTAAACAAAAAACAACAACAACAGCAAT	1500
Db	1441	GGACCTGTATTCCGCAGCAGCGCTTTTAAACAAAAAACAACAACAACAGCAAT	1500
Qy	1501	TTTACCTGGACGTGTGTTCAAAATATAACCTCAATGGGGGTGAATCCATCATCAACCCT	1560
Db	1501	TTTACCTGGACGTGTGTTCAAAATATAACCTCAATGGGGGTGAATCCATCATCAACCCT	1560
Qy	1561	GGCACTGCTATGGCTCACAAAGACGACGAAGACAAGTCTTTCCCATGAGCGGTGTC	1620
Db	1561	GGCACTGCTATGGCTCACAAAGACGACGAAGACAAGTCTTTCCCATGAGCGGTGTC	1620
Qy	1621	ATGATTTTGGAAAAAGAGCGCGGAGCTTCAAAACA CTGCAATGGACAATGTATGAT	1680
Db	1621	ATGATTTTGGAAAAAGAGCGCGGAGCTTCAAAACA CTGCAATGGACAATGTATGAT	1680
Qy	1681	ACAGACGAAGAGGAAATTTAAAGGCACTAAACCTGTGGCCACCGAAAGATTTGGACCGTG	1740
Db	1681	ACAGACGAAGAGGAAATTTAAAGGCACTAAACCTGTGGCCACCGAAAGATTTGGACCGTG	1740
Qy	1741	GCAGTCAATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTCATGCTATGGGA	1800
Db	1741	GCAGTCAATTTCCAGACGACGACAGACCCCTGCGACCGGAGATGTCATGCTATGGGA	1800
Qy	1801	GCATTACCTGGCATGGTGGCAAGATAGACAGTGTACCTGACGGGTCCCATTTGGGCC	1860
Db	1801	GCATTACCTGGCATGGTGGCAAGATAGACAGTGTACCTGACGGGTCCCATTTGGGCC	1860
Qy	1861	AAAAATCTCACA CAGATGGACACTTTTCA CCGGTCTCTCTTATGGCGCGCTTTGGACTC	1920
Db	1861	AAAAATCTCACA CAGATGGACACTTTTCA CCGGTCTCTCTTATGGCGCGCTTTGGACTC	1920
Qy	1921	AAGAACCGCTCTCTCAGATCCTCATCAAAAACA CGCTGCTTCTCGGAATCCTCCGCGG	1980
Db	1921	AAGAACCGCTCTCTCAGATCCTCATCAAAAACA CGCTGCTTCTCGGAATCCTCCGCGG	1980
Qy	1981	GAGTTTTTCAGTCAAAAAGTTTGTCTTCAATCATCACCCAATACTCCA CAGGACAAGTGAGT	2040
Db	1981	GAGTTTTTCAGTCAAAAAGTTTGTCTTCAATCATCACCCAATACTCCA CAGGACAAGTGAGT	2040
Qy	2041	GTGAAAATTTGAATGGGAGCTGCGAAAAAGAAAAA CAGCAAGCGCTGGAAATCCGGAAGTGCAG	2100
Db	2041	GTGAAAATTTGAATGGGAGCTGCGAAAAAGAAAAA CAGCAAGCGCTGGAAATCCGGAAGTGCAG	2100
Qy	2101	TACACATCCCAATATGCAAAATCTGCCAAAGTTGATTTTACTGTGGACAACAATGGACTT	2160
Db	2101	TACACATCCCAATATGCAAAATCTGCCAAAGTTGATTTTACTGTGGACAACAATGGACTT	2160
Qy	2161	TATACTGAGCCTCGCCCATTTGGCACCGCTTACCTTACCCGCTGCCCTGT	2209

[illegible]

Db	1681	ACAGCAGAGAGGAAATTAAGCCACTAAACCTCTGTGGCCACCGAAGATTTTGGGACCGTG	1740	Db	121	GACGGCGGGGTCTGGTCTTCTCGGCTACAAGTACCTCGGACCCCTTCAACGGACTCGAC	180
Qy	1741	GCAGTCAATTTCCAGAGCAGCAGCAGACCTCTGGGACCGGAGATGTGCTATGGGA	1800	Qy	181	AAGGGGAGCCCCGTCAACGCGCGGACGACGCGGCTCTCGAGCAGCAAGGCTTACGAC	240
Db	1741	GCAGTCAATTTCCAGAGCAGCAGCAGACCTCTGGGACCGGAGATGTGCTATGGGA	1800	Db	181	AAGGGGAGCCCCGTCAACGCGCGGACGACGCGGCTCTCGAGCAGCAAGGCTTACGAC	240
Qy	1801	GCATTTACCTGGCATGGTGGCAAGATAGAGACGTGTACTCGAGGGTCCCATTTGGGCC	1860	Qy	241	CAGCAGCTCAAGCGGGTGACAATCCGTACCTGCGGTATTAACACCGCCGACGCCGAGTTT	300
Db	1801	GCATTTACCTGGCATGGTGGCAAGATAGAGACGTGTACTCGAGGGTCCCATTTGGGCC	1860	Db	241	CAGCAGCTCAAGCGGGTGACAATCCGTACCTGCGGTATTAACACCGCCGACGCCGAGTTT	300
Qy	1861	AAAAATTCCTCACACAGATGGACACTTTTCAACCCGCTCTCTCTTATGGGCGGCTTTGGACTC	1920	Qy	301	CAGAGCGCTCTGCAAGAGATACGCTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
Db	1861	AAAAATTCCTCACACAGATGGACACTTTTCAACCCGCTCTCTCTTATGGGCGGCTTTGGACTC	1920	Db	301	CAGAGCGCTCTGCAAGAGATACGCTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
Qy	1921	AAGAACCCGCTCTCAGATGCTCATCAAAAAACAGCGCTGTTCTCGCAATCCTCGGGG	1980	Qy	361	GCCAAAGAGCGGGTCTCGAAACCTCTCGGTCTGGTTGAGGAAGGCGCTTAAAGACGGCTCT	420
Db	1921	AAGAACCCGCTCTCAGATGCTCATCAAAAAACAGCGCTGTTCTCGCAATCCTCGGGG	1980	Db	361	GCCAAAGAGCGGGTCTCGAAACCTCTCGGTCTGGTTGAGGAAGGCGCTTAAAGACGGCTCT	420
Qy	1981	GAGTTTTCAGCTACAAAGTTTCTTCTTATTCATCAACCCATATCTCCACAGGACAAGTGAGT	2040	Qy	421	GGAAAGAAACCTCGGTAGACAGTTCGCAACAGAGCGAGACTCTCTCTCGGGCATCGGC	480
Db	1981	GAGTTTTCAGCTACAAAGTTTCTTCTTATTCATCAACCCATATCTCCACAGGACAAGTGAGT	2040	Db	421	GGAAAGAAACCTCGGTAGACAGTTCGCAACAGAGCGAGACTCTCTCTCGGGCATCGGC	480
Qy	2041	GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGAAATGGCAG	2100	Qy	481	AAGCAGCGCCAGCAGCGCTTAAAAAGAGACTCAATTTTGGTTCAGACTGGGCGACTCAGAG	540
Db	2041	GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCGAAATGGCAG	2100	Db	481	AAGCAGCGCCAGCAGCGCTTAAAAAGAGACTCAATTTTGGTTCAGACTGGGCGACTCAGAG	540
Qy	2101	TACACATCAATTTATGCAAAATCTGCAACAGTGTATTTTATCTGTGGGCAACAATGGACTT	2160	Qy	541	TCAGTCCCGATTCACAACTCTCGGAGAACTCTCAGCAACCTCGGCAACCTCGGTCTGTGGGACCT	600
Db	2101	TACACATCAATTTATGCAAAATCTGCAACAGTGTATTTTATCTGTGTGGCAACAATGGACTT	2160	Db	541	TCAGTCCCGATTCACAACTCTCGGAGAACTCTCAGCAACCTCGGCAACCTCGGTCTGTGGGACCT	600
Qy	2161	TATACAGCTCGCCCATTTGGACCCGTTACCTTACCTTACCCGTCCTGT	2209	Qy	601	ACTCAATGGCTTCAGGCGGTGGCGCAACATGGCAGACATTAACGAAGGCGCCGACCGA	660
Db	2161	TATACAGCTCGCCCATTTGGACCCGTTACCTTACCTTACCCGTCCTGT	2209	Db	601	ACTCAATGGCTTCAGGCGGTGGCGCAACATGGCAGACATTAACGAAGGCGCCGACCGA	660
RESULT 3				Qy	661	GTGGTAAATGCTCAGGAAATTTGGCATTTGCATTTCCATATGCGTGGGCGACAGAGTCAATC	720
LOCUS	CS073614	2211 bp	DNA	linear	PAT 05-MAY-2005		
DEFINITION	Sequence 202 from Patent WO200503321.			Db	661	GTGGTAAATGCTCAGGAAATTTGGCATTTGCATTTCCATATGCGTGGGCGACAGAGTCAATC	720
ACCESSION	CS073614			Qy	721	ACCACGACGCCGACCTGGGCTTGGCCACCTTACAAATTAACCACTTACAGGAATTC	780
VERSION	CS073614.1	GI:63090494		Db	721	ACCACGACGCCGACCTGGGCTTGGCCACCTTACAAATTAACCACTTACAGGAATTC	780
KEYWORDS	unidentified			Qy	781	TCCAGTGTCTCAACGGGGCCAGCAACGACCACTACTTCCGCTACAGCACCCCTGG	840
SOURCE	unidentified			Db	781	TCCAGTGTCTCAACGGGGCCAGCAACGACCACTACTTCCGCTACAGCACCCCTGG	840
ORGANISM	unclassified.			Qy	841	GGGTATTTGATTTCAACAGATTTCCACTGCGCACTTTTCAACACGTCGTCGCGAGCTC	900
REFERENCE	1	Wilson, J.M., Gao, G., Alvira, M.R. and Vandenberghe, L.H.		Db	841	GGGTATTTGATTTCAACAGATTTCCACTGCGCACTTTTCAACACGTCGTCGCGAGCTC	900
AUTHORS		Adeno-associated virus (aav) clades, sequences, vectors containing		Qy	901	ATCAACAAATTTGGGATTCGGGCCCAAGAGACTCAACTTCAAACTTTTCAACATCCAA	960
TITLE		same, and uses therefor		Db	901	ATCAACAAATTTGGGATTCGGGCCCAAGAGACTCAACTTCAAACTTTTCAACATCCAA	960
JOURNAL		Patent: WO 200503321-A 202 14-APR-2005;		Qy	961	GTCAAGGAGTTCACGACGAATGATGGCGTCAACCATCGCTTAATACTTACAGGACG	1020
		The Trustees of the University of Pennsylvania (US)		Db	961	GTCAAGGAGTTCACGACGAATGATGGCGTCAACCATCGCTTAATACTTACAGGACG	1020
FEATURES		Location/Qualifiers		Qy	1021	GTTCAGTCTTCTCGGATTCGAGTACCAAGTTCGCTTCCGTCGCTTCCGTCGACCCAG	1080
source		1. .2211		Db	1021	GTTCAGTCTTCTCGGATTCGAGTACCAAGTTCGCTTCCGTCGCTTCCGTCGACCCAG	1080
		/organism="unidentified"		Qy	1081	GGCTGCTTCTTCCGCGGACGCTGTTTCAATGATTCGCAATACGGCTTACCTGAGC	1140
		/mol_type="unassigned DNA"		Db	1081	GGCTGCTTCTTCCGCGGACGCTGTTTCAATGATTCGCAATACGGCTTACCTGAGC	1140
		/db_xref="taxon:32644"		Qy	1141	CTCAACAAATGGCAGCAAGCGGTGGAGCTTTCATCTTCTTACTGCTGGAAATTTTCCCT	1200
		/note="adeno-associated virus, serotype 1"		Db	1141	CTCAACAAATGGCAGCAAGCGGTGGAGCTTTCATCTTCTTACTGCTGGAAATTTTCCCT	1200
ORIGIN				Qy	1201	TCCTCAGATGCTGAGAACCGGCAACAACTTTTACCTTTCAGCTACACCTTTGAGGAAGTGCCT	1260
Query Match		100.0%; Score 2209; DB 6; Length 2211;		Db	1201	TCCTCAGATGCTGAGAACCGGCAACAACTTTTACCTTTCAGCTACACCTTTGAGGAAGTGCCT	1260
Best Local Similarity		100.0%; Pred. No. 0;		Qy			
Matches 2209; Conservative		0; Mismatches 0; Indels 0; Gaps 0;		Db			
Qy	1	ATGGCTGCCGATGGTTATCTTCAGATTTGGCTCGAGGACAACTCTCTGAGGCGATTCCG	60	Qy	1081	GGCTGCTTCTTCCGCGGACGCTGTTTCAATGATTCGCAATACGGCTTACCTGAGC	1140
Db	1	ATGGCTGCCGATGGTTATCTTCAGATTTGGCTCGAGGACAACTCTCTGAGGCGATTCCG	60	Db	1081	GGCTGCTTCTTCCGCGGACGCTGTTTCAATGATTCGCAATACGGCTTACCTGAGC	1140
Qy	61	GAGTGTGGGACTTGAAACCTTGAGCCCGCAAGCCCAAGAGCCCAACAGCAAAAGCAGGAC	120	Qy	1141	CTCAACAAATGGCAGCAAGCGGTGGAGCTTTCATCTTCTTACTGCTGGAAATTTTCCCT	1200
Db	61	GAGTGTGGGACTTGAAACCTTGAGCCCGCAAGCCCAAGAGCCCAACAGCAAAAGCAGGAC	120	Db	1141	CTCAACAAATGGCAGCAAGCGGTGGAGCTTTCATCTTCTTACTGCTGGAAATTTTCCCT	1200
Qy	121	GACGGCGGGTCTGGTCTTCTCGGCTACAGTACCTCGGACCCCTTCAACGGACTCGAC	180	Qy	1201	TCCTCAGATGCTGAGAACCGGCAACAACTTTTACCTTTCAGCTACACCTTTGAGGAAGTGCCT	1260
Db				Db	1201	TCCTCAGATGCTGAGAACCGGCAACAACTTTTACCTTTCAGCTACACCTTTGAGGAAGTGCCT	1260

```
QY 1261 TTCCACAGAGCTACGCGCACAGCCAGAGCCTGGACCGCGTGATGAATCCTCTCATCGAC 1320
|
|
|
Db 1261 TTCCACAGAGCTACGCGCACAGCCAGAGCCTGGACCGCGTGATGAATCCTCTCATCGAC 1320
|
|
|
QY 1321 CAATACCTGTATTACCTGACAGAACTCAAAATCAGTCCGGAGTCCCAAAACAGGAC 1380
|
|
|
Db 1321 CAATACCTGTATTACCTGACAGAACTCAAAATCAGTCCGGAGTCCCAAAACAGGAC 1380
|
|
|
QY 1381 TTGCTGTTTACGCGTGGTCTCCAGCTGGCATGCTCTTTCAGCCCAAAACTGGCTACCT 1440
|
|
|
Db 1381 TTGCTGTTTACGCGTGGTCTCCAGCTGGCATGCTCTTTCAGCCCAAAACTGGCTACCT 1440
|
|
|
QY 1441 GGACCTGTTTATGGCAGCAGCGCGTTTCTAAACAAACACAGACAAACACAGCAAT 1500
|
|
|
Db 1441 GGACCTGTTTATGGCAGCAGCGCGTTTCTAAACAAACACAGACAAACACAGCAAT 1500
|
|
|
QY 1501 TTTACCTGGAGCTGGTCTTCAAAATATAACCTCAATGGGGGTGAATCCATCATCAACCT 1560
|
|
|
Db 1501 TTTACCTGGAGCTGGTCTTCAAAATATAACCTCAATGGGGGTGAATCCATCATCAACCT 1560
|
|
|
QY 1561 GGCACCTGCTATGSCCTCACACAAAGACAGCAAGTCTTTCCTCATGAGCGGTGTC 1620
|
|
|
Db 1561 GGCACCTGCTATGSCCTCACACAAAGACAGCAAGTCTTTCCTCATGAGCGGTGTC 1620
|
|
|
QY 1621 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAACACACTGCAATTTGGACAAATGTATGAT 1680
|
|
|
Db 1621 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAACACACTGCAATTTGGACAAATGTATGAT 1680
|
|
|
QY 1681 ACAGACGAAGAGAAATTAAGCCACTAACCTTGTGGCCACCGAAAGATTTGGGACCGTG 1740
|
|
|
Db 1681 ACAGACGAAGAGAAATTAAGCCACTAACCTTGTGGCCACCGAAAGATTTGGGACCGTG 1740
|
|
|
QY 1741 CGAGTCAATTTCCAGAGCAGCAGCAGACAGCCCTGCGACCGAGATGTGATGCTATGGGA 1800
|
|
|
Db 1741 CGAGTCAATTTCCAGAGCAGCAGCAGACAGCCCTGCGACCGAGATGTGATGCTATGGGA 1800
|
|
|
QY 1801 GCATTACCTGGCATGTTGGCAAGATAGAGACGTGTACCTGCAAGGTCCCATTTGGGCC 1860
|
|
|
Db 1801 GCATTACCTGGCATGTTGGCAAGATAGAGACGTGTACCTGCAAGGTCCCATTTGGGCC 1860
|
|
|
QY 1861 AAAATTTCTCACACAGATGACACTTTTCAACCGCTCTCTTATGGGGCGCTTTGACTC 1920
|
|
|
Db 1861 AAAATTTCTCACACAGATGACACTTTTCAACCGCTCTCTTATGGGGCGCTTTGACTC 1920
|
|
|
QY 1921 AAGAACCCGCTCTCAGATCCTCATCAAAACACCGCTGTTCTCGAATCCTCCGGCG 1980
|
|
|
Db 1921 AAGAACCCGCTCTCAGATCCTCATCAAAACACCGCTGTTCTCGAATCCTCCGGCG 1980
|
|
|
QY 1981 GAGTTTTCAGCTACAAAGTTTGCTTCATTTCATCAACCAATACCTCCACAGGACAAGTGAGT 2040
|
|
|
Db 1981 GAGTTTTCAGCTACAAAGTTTGCTTCATTTCATCAACCAATACCTCCACAGGACAAGTGAGT 2040
|
|
|
QY 2041 GTGGAATTAATGGAGTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCGGAAGTGCAG 2100
|
|
|
Db 2041 GTGGAATTAATGGAGTGCAGAAAGAAAACAGCAAGCGCTGGAAATCCGGAAGTGCAG 2100
|
|
|
QY 2101 TACACATCAATTAATGCAAAATCTGCCAAGTTGATTTTACTTGGACAAATGCACTT 2160
|
|
|
Db 2101 TACACATCAATTAATGCAAAATCTGCCAAGTTGATTTTACTTGGACAAATGCACTT 2160
|
|
|
QY 2161 TATACTGAGCCTCGCCCCATTTGGCACCCTGTTACTTACCCGTCCCTGT 2209
|
|
|
Db 2161 TATACTGAGCCTCGCCCCATTTGGCACCCTGTTACTTACCCGTCCCTGT 2209
|
|
|
```

```
RESULT 4
AR562503 LOCUS 2211 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 12 from patent US 6759237.
ACCESSION AR562503
VERSION AR562503.1 GI:53976569
KEYWORDS
```

```
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 2211)
TITLE Wilson,J.M. and Xiao,W.
JOURNAL Adeno-associated virus serotype 1 nucleic acid sequences, vectors
and host cells containing same
Patent: US 6759237-A 12 06-JUL-2004;
The Trustees of the University of Pennsylvania; Philadelphia, PA
FEATURES
Location/Qualifiers
1..2211
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 2209; DB 6; Length 2211;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGTGGCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCG 60
|
|
|
Db 1 ATGGTGGCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCG 60
|
|
|
QY 61 GAGTGGTGGACTTGAACCTTGGAGCCCGGAGCCCAAGCCAAACAGCAAGCAGGAC 120
|
|
|
Db 61 GAGTGGTGGACTTGAACCTTGGAGCCCGGAGCCCAAGCCAAACAGCAGGAGGAC 120
|
|
|
QY 121 GACGGCCGGGCTGTGGTGTCTTCTGGCTACAAAGTACTCTCGGACCTCTTCAACGAGCTCGAC 180
|
|
|
Db 121 GACGGCCGGGCTGTGGTGTCTTCTGGCTACAAAGTACTCTCGGACCTCTTCAACGAGCTCGAC 180
|
|
|
QY 181 AAGGGGAGCCCCGTCAACCGCGGACGCGCGGCTCGAGCAAGAGCCTTACGAC 240
|
|
|
Db 181 AAGGGGAGCCCCGTCAACCGCGGACGCGCGGCTCGAGCAAGAGCCTTACGAC 240
|
|
|
QY 241 CAGCAGCTCAAGCGGCTGACAACTCGGTACCTCGGTATAAACACGCGCGCGAGTTT 300
|
|
|
Db 241 CAGCAGCTCAAGCGGCTGACAACTCGGTATCGGTATAAACACGCGCGCGAGTTT 300
|
|
|
QY 301 CAGGAGCGCTCGCAAGAAGATACGTCTTTTGGGGCAACCTCTGGGCGAGCAGTCTTCCAG 360
|
|
|
Db 301 CAGGAGCGCTCGCAAGAAGATACGTCTTTTGGGGCAACCTCTGGGCGAGCAGTCTTCCAG 360
|
|
|
QY 361 GCCAAGAAGCGGTTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTAAGCGGTCT 420
|
|
|
Db 361 GCCAAGAAGCGGTTTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTAAGCGGTCT 420
|
|
|
QY 421 GGAAGAAGCGTCCGGTAGAGCAGTCCCAAGAGCCAGACTCTCTCGGGCATCGGC 480
|
|
|
Db 421 GGAAGAAGCGTCCGGTAGAGCAGTCCCAAGAGCCAGACTCTCTCGGGCATCGGC 480
|
|
|
QY 481 AAGACAGGCGCAGCGCCCTGTAAGAGAGACTCAATTTTGGTTCAGACTTGGGACTCAGAG 540
|
|
|
Db 481 AAGACAGGCGCAGCGCCCTGTAAGAGAGACTCAATTTTGGTTCAGACTTGGGACTCAGAG 540
|
|
|
QY 541 TCAGTCCCCGATCCCAACCTCTCGGAGAACTCTCGAGCAACCCCGCTCTCTGGGACCT 600
|
|
|
Db 541 TCAGTCCCCGATCCCAACCTCTCGGAGAACTCTCGAGCAACCCCGCTCTCTGGGACCT 600
|
|
|
QY 601 ACTAATATGGCTTCAGCGGTGGCGCACCAATGGGAGACAATAACGAAGCGCGGACGGA 660
|
|
|
Db 601 ACTAATATGGCTTCAGCGGTGGCGCACCAATGGGAGACAATAACGAAGCGCGGACGGA 660
|
|
|
QY 661 GTGGGTAATGCCTCAGGAAATTTGGCATTTCCATGGCTGGGCGACAGAGTCATC 720
|
|
|
Db 661 GTGGGTAATGCCTCAGGAAATTTGGCATTTCCATGGCTGGGCGACAGAGTCATC 720
|
|
|
QY 721 ACCACGACGACCCGCACTTGGGCTTGGCCCACTCAATAAACCACTCTTACAGCAAAATC 780
|
|
|
Db 721 ACCACGACGACCCGCACTTGGGCTTGGCCCACTCAATAAACCACTCTTACAGCAAAATC 780
|
|
|
QY 781 TCCAGTGTCTTCAAGCGGGGCGCAGCAACGCAACCACTACTTGGGCTACAGCACCCCTGG 840
|
|
|
```

Db 781 TCCAGTGTCTTCAACGGGGCCAGCAACGACCAACCACTACTTCGGCTACAGCACCCCTGG 840  
QY 841 GGGTATTTTGAATTTCAACAGATTCACCTGCCACTTTTTCACCACTGAGTGGCAGGACTC 900  
Db 841 GGGTATTTTGAATTTCAACAGATTCACCTGCCACTTTTTCACCACTGAGTGGCAGGACTC 900  
QY 901 ATCAACAACAATTTGGGATTCGGCCCAAGAGACTCAAACTTCAAACTCTTCAACATCCAA 960  
Db 901 ATCAACAACAATTTGGGATTCGGCCCAAGAGACTCAAACTTCAAACTCTTCAACATCCAA 960  
QY 961 GTCAAGAGAGTCAAGACGAATGATGGCGTCACAACCATCGCTAAATAACCTTACCAGCAG 1020  
Db 961 GTCAAGAGAGTCAAGACGAATGATGGCGTCACAACCATCGCTAAATAACCTTACCAGCAG 1020  
QY 1021 GTTCAAGTCTTTCGAGCTCGAGTACCAAGCTTTCGATCGCTCTCGGCTCTCGCACCCAG 1080  
Db 1021 GTTCAAGTCTTTCGAGCTCGAGTACCAAGCTTTCGATCGCTCTCGGCTCTCGCACCCAG 1080  
QY 1081 GGTGCTCTCCCTCCGTTCCGGGGGAGCGTTTCATGATTCGGCAATACGGCTACCTGAGC 1140  
Db 1081 GGTGCTCTCCCTCCGTTCCGGGGGAGCGTTTCATGATTCGGCAATACGGCTACCTGAGC 1140  
QY 1141 CTCAACAATGGCAGCAAGCGTGGGACGTTTCATCTCTTTTACTGCTGGAATATTTCCT 1200  
Db 1141 CTCAACAATGGCAGCAAGCGTGGGACGTTTCATCTCTTTTACTGCTGGAATATTTCCT 1200  
QY 1201 TCTCAGATCTGAGAACGGGCAACAATTTTACCTTCAGTACACCTTTGAGGAAGTGCT 1260  
Db 1201 TCTCAGATCTGAGAACGGGCAACAATTTTACCTTCAGTACACCTTTGAGGAAGTGCT 1260  
QY 1261 TTCCAGCAGCTACGCGCACAGCCAGCGCTGGACCGCTGATGATCTCTCATCGAC 1320  
Db 1261 TTCCAGCAGCTACGCGCACAGCCAGCGCTGGACCGCTGATGATCTCTCATCGAC 1320  
QY 1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAACCAAGGAC 1380  
Db 1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCCGGAAGTCCCAACCAAGGAC 1380  
QY 1381 TTGCTCTTTAGCGCTGGGTCTCCAGCTGGCATGTCTGTTTTCAGCCCAAAAACCTGGCT 1440  
Db 1381 TTGCTCTTTAGCGCTGGGTCTCCAGCTGGCATGTCTGTTTTCAGCCCAAAAACCTGGCT 1440  
QY 1441 GGAACCTGTATTCCGCGAGCGCGTCTTAAACCAAAAACAGACAGCAACAGCAAT 1500  
Db 1441 GGAACCTGTATTCCGCGAGCGCGTCTTAAACCAAAAACAGACAGCAACAGCAAT 1500  
QY 1501 TTTACCTGAGCTGGTGCTTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCT 1560  
Db 1501 TTTACCTGAGCTGGTGCTTCAAAATATAACCTCAATGGCGTGAATCCATCATCAACCT 1560  
QY 1561 GGCACCTGTATGGCTTCACAAAGACGACGAAAGTTCCTTTCCCATGAGCGGTGTC 1620  
Db 1561 GGCACCTGTATGGCTTCACAAAGACGACGAAAGTTCCTTTCCCATGAGCGGTGTC 1620  
QY 1621 ATGATTTTGGAAAAAGAGCGCGGAGCTTCAACACATGCTGATTCGACATGCTGATTT 1680  
Db 1621 ATGATTTTGGAAAAAGAGCGCGGAGCTTCAACACATGCTGATTCGACATGCTGATTT 1680  
QY 1681 ACAGACGAAGAGAAATTAAGGCACTTAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740  
Db 1681 ACAGACGAAGAGAAATTAAGGCACTTAACCTGTGGCCACCGAAAGATTTGGGACCGTG 1740  
QY 1741 GCAGTCAATTTTCAGAGCAGCAGACAGACCTTCGACCGGAGATGTGATGCTATGGGA 1800  
Db 1741 GCAGTCAATTTTCAGAGCAGCAGACAGACCTTCGACCGGAGATGTGATGCTATGGGA 1800  
QY 1801 GCATACCTGGCATGTGTGGCAAGATAGAGAGCTGTACTGACGGTCCCATTTGGGC 1860  
Db 1801 GCATACCTGGCATGTGTGGCAAGATAGAGAGCTGTACTGACGGTCCCATTTGGGC 1860  
QY 1861 AAAATTCCTCACACAGATGACACTTTTACCCCTCTCTTATGGCGGCTTTGGACTC 1920  
Db 1861 AAAATTCCTCACACAGATGACACTTTTACCCCTCTCTTATGGCGGCTTTGGACTC 1920

QY 1921 AAGAACCGCCTCTCTCAGATCTCTATCAAAAACAGCCTGTTCCTGCGAATCCTCCGGCG 1980  
Db 1921 AAGAACCGCCTCTCTCAGATCTCTATCAAAAACAGCCTGTTCCTGCGAATCCTCCGGCG 1980  
QY 1981 GAGTTTTTCAGCTACAAAGTTTGTCTTCAATTCATCACCCTCAATCTCCACAGGCAAGTGAGT 2040  
Db 1981 GAGTTTTTCAGCTACAAAGTTTGTCTTCAATTCATCACCCTCAATCTCCACAGGCAAGTGAGT 2040  
QY 2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGTGGAAATCCCGAAGTGAG 2100  
Db 2041 GTGGAAATTTGAATGGGAGCTGCAGAAAGAAACAGCAAGCGTGGAAATCCCGAAGTGAG 2100  
QY 2101 TACACATCAATATGCAAAATCTGCCCAACGTTGATTTTACTGTGGACAAATGGACTT 2160  
Db 2101 TACACATCAATATGCAAAATCTGCCCAACGTTGATTTTACTGTGGACAAATGGACTT 2160  
QY 2161 TATACGTAGCCTCGCCCATTTGGCACCCGTTACCTTACCCGTCCTCTGT 2209  
Db 2161 TATACGTAGCCTCGCCCATTTGGCACCCGTTACCTTACCCGTCCTCTGT 2209

RESULT 5  
BD242766  
LOCUS BD242766  
DEFINITION Adeno-associated virus serum type 1 nucleic acid sequence, vector and host cell containing the same.  
ACCESSION BD242766  
VERSION BD242766.1 GI:33052536  
KEYWORDS JP 2002529098-A/1.  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 4718)  
AUTHORS Wilson, J.M. and Xiao, W.  
TITLE Adeno-associated virus serum type 1 nucleic acid sequence, vector and host cell containing the same  
JOURNAL Patent: JP 2002529098-A 1 10-SEP-2002;  
THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA  
COMMENT OS AAV-1  
PN JP 2002529098-A/1  
PD 10-SEP-2002  
PF 02-NOV-1999 JP 2000581227  
PR 05-NOV-1998 US 60/107114  
PI JAMES M WILSON, WEIDONG XIAO  
PC C12N15/09, A61K31/711, A61K48/00, A61P43/00, C12N1/15, C12N1/19, PC C12N1/21.  
PC C12N5/10, C12P21/02//A61K35/76, C12N15/00, C12N5/00 CC  
Adeno-associated virus serum type 1 nucleic acid sequence, CC  
vector and host  
CC cell containing the same  
FH Key Location/Qualifiers  
FT CDS (335) .. (2206)  
FT CDS (2223) .. (4430).  
FEATURES  
source  
1..4718  
/organism="unidentified"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

ORIGIN  
Query Match 100.0%; Score 2209; DB 6; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTGCGGATGGTATCTTCCAGATTGGCTCGAGGCAACCTCTCTGAGGGCATTCGC 60  
Db 2223 ATGGCTGCGGATGGTATCTTCCAGATTGGCTCGAGGCAACCTCTCTGAGGGCATTCGC 2282  
QY 61 GAGTGTGGGACTTGAACCTTGAGCCCGGAGCCCAAGCCCAAGCAACAGCAAGCAGGAC 120  
Db 2283 GAGTGTGGGACTTGAACCTTGAGCCCGGAGCCCAAGCCCAAGCAACAGCAGGAC 2342



QY 121 GACGCGCGGGTCTGGTCTTCTTGCTACAGTACCTCGGACCTTCAACGACTCGAC 180  
DB 2343 GACGCGCGGGTCTGGTCTTCTTGCTACAGTACCTCGGACCTTCAACGACTCGAC 2402  
QY 181 AAGGGGAGCCGCTCAACCGGGCGGACGAGCGGCCCTCGAGCA CGAAGAAGGCTACGAC 240  
DB 2403 AAGGGGAGCCGCTCAACCGGGCGGACGAGCGGCCCTCGAGCA CGAAGAAGGCTACGAC 2462  
QY 241 GACGAGCTCAAAAGCGGGTCGAATCCGTACTCGGTATATAA CACGCGCGACGCCGAGTTT 300  
DB 2463 GACGAGCTCAAAAGCGGGTGCAATCCGTACTCGGTATATAA CACGCGCGACGCCGAGTTT 2522  
QY 301 CAGGAGCGCTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360  
DB 2523 CAGGAGCGCTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 2582  
QY 361 GCCAAGAAGCGGGTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT 420  
DB 2583 GCCAAGAAGCGGGTCTCGAACTCTCGGTCTGGTTGAGGAAGCGCTAAGACGGCTCCT 2642  
QY 421 GGAAGAAGACGTCGGGTAGAGCAGTCGCCACAAGACCGAGACTCTCTCGGSCATCGGC 480  
DB 2643 GGAAGAAGACGTCGGGTAGAGCAGTCGCCACAAGACCGAGACTCTCTCGGSCATCGGC 2702  
QY 481 AAGACAGGCGCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTCGCGACTCAGAG 540  
DB 2703 AAGACAGGCGCAGCAGCCCGCTAAAGAGAGACTCAATTTTGGTCAGACTCGCGACTCAGAG 2762  
QY 541 TCAGTCCCCGATCCAACTCTTCGGAGAACTTCCAGCAACCCCCCGCTGCTGTGGACCT 600  
DB 2763 TCAGTCCCCGATCCAACTCTTCGGAGAACTTCCAGCAACCCCCCGCTGCTGTGGACCT 2822  
QY 601 ACTACAAATGGCTTCAGCGGTGCGCACCAATGGCAGACAATAA CCAAGCGCGCGACGGA 660  
DB 2823 ACTACAAATGGCTTCAGCGGTGCGCACCAATGGCAGACAATAA CCAAGCGCGCGACGGA 2882  
QY 661 GTGGGTAAATGGCTTCAGGAAATGGCAATTCGATTCGATTC CACATGGTGGCGCAGAGTCATC 720  
DB 2883 GTGGGTAAATGGCTTCAGGAAATGGCAATTCGATTCGATTC CACATGGTGGCGCAGAGTCATC 2942  
QY 721 ACCACAGCACCCGCACTGGGCTTGGCCCACTCAATAACCACTT CATAAGCAATC 780  
DB 2943 ACCACAGCACCCGCACTGGGCTTGGCCCACTCAATAACCACTT CATAAGCAATC 3002  
QY 781 TCCAGTGTCTCAACGGGGCCAGCAACGCAACCACTACTTCGG CTACAGCACCCCTGG 840  
DB 3003 TCCAGTGTCTCAACGGGGCCAGCAACGCAACCACTACTTC GGCTACAGCACCCCTGG 3062  
QY 841 GGGTATTTTGTATTTCAACAGATTTCCACTGCCACTTTTCA CACAGTGACTGGCAGCGACTC 900  
DB 3063 GGGTATTTTGTATTTCAACAGATTTCCACTGCCACTTTT CACCACTGACTGGCAGCGACTC 3122  
QY 901 ATCAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTTC CAACTCTTCAACATCCAA 960  
DB 3123 ATCAACAAATTTGGGATTTCCGGCCCAAGAGACTCAACTT CAACTCTTCAACATCCAA 3182  
QY 961 GTCAAGAGGTACGCAAGATGAGCTCACAACCACTGCTAAT AACTTACCTTACGACAG 1020  
DB 3183 GTCAAGAGGTACGCAAGATGAGCTCACAACCACTGCTAAT AACTTACCTTACGACAG 3242  
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCGTAAGT CCGTCTCGGCTCTGGCAGCCAG 1080  
DB 3243 GTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCGTAAGT CCGTCTCGGCTCTGGCAGCCAG 3302  
QY 1081 GGCTGCTCTCTCGGTTCCGGGGAGCGTGTTCATGATTCG GAAATACGGCTACCTGAGC 1140  
DB 3303 GGCTGCTCTCTCGGTTCCGGGGAGCGTGTTCATGATTCG GAAATACGGCTACCTGAGC 3362  
QY 1141 CTCAACAAATGGCAGCCAGCGTGGGAGCTTCATCTTTTACT GCGCTGGAATATTTCCCT 1200  
DB 3363 CTCAACAAATGGCAGCCAGCGTGGGAGCTTCATCTTTTACT GCGCTGGAATATTTCCCT 3422  
QY 1201 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTCA GCTACACCTTTGAGGAAGTGCCT 1260

DB 3423 TCTCAGATGCTGAGAACGGGCAACAACTTTTACCTTCA GCTACACCTTTGAGGAAGTGCCT 3482  
QY 1261 TTCCACAGCAGCTACGCGCACAGCCGAGCCTTGGACCG GCTCATGAATCTCTCTCATCGAC 1320  
DB 3483 TTCCACAGCAGCTACGCGCACAGCCGAGCCTTGGACCG GCTCATGAATCTCTCTCATCGAC 3542  
QY 1321 CAATACCTGTATTACCTGAACAGAACTCAAAATCA GTCCGGAAGTCCCAAAACAAGGAC 1380  
DB 3543 CAATACCTGTATTACCTGAACAGAACTCAAAATCA GTCCGGAAGTCCCAAAACAAGGAC 3602  
QY 1381 TTGCTGTTTACGCGTGGTCTCCAGCTGGCATGTCTGT TACGCCCAAAACCTGGCTACCT 1440  
DB 3603 TTGCTGTTTACGCGTGGTCTCCAGCTGGCATGTCT GTTACGCCCAAAACCTGGCTACCT 3662  
QY 1441 GGACCTGTATTACGCGCAGCAGCGCTTTCTAAAAA CAAAAAAGAGAACTCAAAACAAGCAAT 1500  
DB 3663 GGACCTGTATTACGCGCAGCAGCGCTTTCTAAAA CAAAAAAGAGAACTCAAAACAAGCAAT 3722  
QY 1501 TTTCCTGGACTGGTCTTCAAAATATAA CTTCAATGGCGGTGAATCCATCATCAACCTT 1560  
DB 3723 TTTCCTGGACTGGTCTTCAAAATATAA CTTCAATGGCGGTGAATCCATCATCAACCTT 3782  
QY 1561 GGCACCTGCTATGGCTTCAACAAAGACGACGAAGACA GTTCTTTCCCATGAGCGGTTC 1620  
DB 3783 GGCACCTGCTATGGCTTCAACAAAGACGACGAAGACA GTTCTTTCCCATGAGCGGTTC 3842  
QY 1621 ATGATTTTGGAAAAAGAGAGCGCCGAGCTTCAAA CACTTGGACAAATGTCTATGAT 1680  
DB 3843 ATGATTTTGGAAAAAGAGAGCGCCGAGCTTCAAA CACTTGGACAAATGTCTATGAT 3902  
QY 1681 ACAGACGAAGAGAGAAATTAAGGCCACTAA CCGCTGTGGCCACCGAAAGATTTGGGACCGTG 1740  
DB 3903 ACAGACGAAGAGAGAAATTAAGGCCACTAA CCGCTGTGGCCACCGAAAGATTTGGGACCGTG 3962  
QY 1741 GAGTCAATTTCCAGAGCAGCAGCAGACCCCTGGAC CCGGAGATGTGATGTATGSSGA 1800  
DB 3963 GCAGTCAATTTCCAGAGCAGCAGCAGACCCCTGG ACACCGGAGATGTGATGTATGSSGA 4022  
QY 1801 GCATTTACCTGGCATGTGTGGCAAGATAGAGACGT GTACTGCAGGGTCTCCATTTGGGCC 1860  
DB 4023 GCATTTACCTGGCATGTGTGTGGCAAGATAGAGACGT GTACTGCAGGGTCTCCATTTGGGCC 4082  
QY 1861 AAAATTTCTCACACAGATGACACTTTTCA CCGCTCTCTCTTATAGGGCGCTTTGACTC 1920  
DB 4083 AAAATTTCTCACACAGATGACACTTTTCA CCGCTCTCTCTTATAGGGCGCTTTGACTC 4142  
QY 1921 AAGAACCCGCTCTCAGATCTCATCAAAAACA CCGCTGTCTCTGGGATCTCTCCGGCG 1980  
DB 4143 AAGAACCCGCTCTCAGATCTCATCAAAAACA CCGCTGTCTCTGGGATCTCTCCGGCG 4202  
QY 1981 GAGTTTTCAGCTACAAAGTTTGGCTTCATTCAC CCAATACTCCACAGGACAAGTGAGT 2040  
DB 4203 GAGTTTTCAGCTACAAAGTTTGGCTTCATTCAC CCAATACTCCACAGGACAAGTGAGT 4262  
QY 2041 GTGGAAATTTGAATGGGAGCTGCAAGAAAGAAA CAGCAAGCGCTGGAATCCCGAAGTGCAG 2100  
DB 4263 GTGGAAATTTGAATGGGAGCTGCAAGAAAGAAA CAGCAAGCGCTGGAATCCCGAAGTGCAG 4322  
QY 2101 TACACATCCAAATTTATGCAAAAATCTGCCAAAG TTTGATTTTACTGTGACAAATGACAT 2160  
DB 4323 TACACATCCAAATTTATGCAAAAATCTGCCAAAG TTTGATTTTACTGTGACAAATGACAT 4382  
QY 2161 TATACCTGAGCTCGCCCATTTGGCACCCGTTACTT TACCGTCCCTCTGT 2209  
DB 4383 TATACCTGAGCTCGCCCATTTGGCACCCGTTACTT TACCGTCCCTCTGT 4431

RESULT 6

AR562498

LOCUS

DEFINITION

ACCESSION

AR562498

Sequence 1 from patent US 6759237.

AR562498

4718 bp

DNA

linear

PAT 08-OCT-2004



AR562498.1 GI:53976564  
VERSION  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 4718)  
AUTHORS Wilson, J.M. and Xiao, W.  
TITLE Adeno-associated virus serotype 1 nucleic acid sequences, vectors and host cells containing same  
JOURNAL Patent: US 6759237-A 1 06-JUL-2004;  
The Trustees of the University of Pennsylvania; Philadelphia, PA  
FEATURES  
Location/Qualifiers  
1..4718  
source /organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 100.0%; Score 2209; DB 6; Length 4718;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGCTGCCGATGGTATCTTCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC 60  
Db 2223 ATGGCTGCCGATGGTATCTTCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGC 2282  
QY 61 GAGTGGTGGGACTTGAAACCTGGAGCCCGAAGCCCAAGCCCAAGCCCAAGCAAGCAGGAC 120  
Db 2283 GAGTGGTGGGACTTGAAACCTGGAGCCCGAAGCCCAAGCCCAAGCCCAAGCAAGCAGGAC 2342  
QY 121 GACGGCCGGGCTCGGTGCTTCTGGCTACAACTACCTCGGACCCCTTCAACGACTCGAC 180  
Db 2343 GACGGCCGGGCTCGGTGCTTCTGGCTACAACTACCTCGGACCCCTTCAACGACTCGAC 2402  
QY 181 AAGGGGAGCCCGTCAACCGCGGACGACGCGGCCCTCGAGCAGCAAGGCCCTTACGAC 240  
Db 2403 AAGGGGAGCCCGTCAACCGCGGACGACGCGGCCCTCGAGCAGCAAGGCCCTTACGAC 2462  
QY 241 CAGCAGCTCAAGCGGTGACAACTCGTACCTCGGTATACCAAGCCGACGCGGAGTTT 300  
Db 2463 CAGCAGCTCAAGCGGTGACAACTCGTACCTCGGTATACCAAGCCGACGCGGAGTTT 2522  
QY 301 CAGGAGCGTCTCAAGAGATAGCTCTTTTGGGGCAACCTCGGGGAGCAGTCTTCCAG 360  
Db 2523 CAGGAGCGTCTCAAGAGATAGCTCTTTTGGGGCAACCTCGGGGAGCAGTCTTCCAG 2582  
QY 361 GCCAAGAGCGGGTCTCGAACTCTCGTCTGGTTGAGGAAGGCGCTTAAGACGGCTCCT 420  
Db 2583 GCCAAGAGCGGGTCTCGAACTCTCGTCTGGTTGAGGAAGGCGCTTAAGACGGCTCCT 2642  
QY 421 GGAAAGAAACGTCGGTAGCAGTCCGACAGAGCCGAGACTCTCTCGGGCATCGGC 480  
Db 2643 GGAAAGAAACGTCGGTAGCAGTCCGACAGAGCCGAGACTCTCTCGGGCATCGGC 2702  
QY 481 AAGACAGGCGAGCAGCCCGCTAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 540  
Db 2703 AAGACAGGCGAGCAGCCCGCTAAAAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAG 2762  
QY 541 TCAGTCCCAGATCCCAACCTCTCGGAGAACTTCCAGCAACCCCGCTCGTGGGACCT 600  
Db 2763 TCAGTCCCAGATCCCAACCTCTCGGAGAACTTCCAGCAACCCCGCTCGTGGGACCT 2822  
QY 601 ACTACAATGGCTTCAGCGGTGGCGCAACCAATGGGAGCAATAAAGCAAGCGCCGACGGA 660  
Db 2823 ACTACAATGGCTTCAGCGGTGGCGCAACCAATGGGAGCAATAAAGCAAGCGCCGACGGA 2882  
QY 661 GTGGGTAATGCTTCAGGAAATGGCAATTCGATTTCACATGGCTGGCGCAGAGTCAATC 720  
Db 2883 GTGGGTAATGCTTCAGGAAATGGCAATTCGATTTCACATGGCTGGCGCAGAGTCAATC 2942  
QY 721 ACCACGAGCAACCGCACTTGGGCTTGGCCCACTCAATAACCACTCTTACAGCAAAATC 780  
Db 2943 ACCACGAGCAACCGCACTTGGGCTTGGCCCACTCAATAACCACTCTTACAGCAAAATC 3002

QY 781 TCCAGTGCTTCAACGGGGGCCAGCAACGACCAACCACTACTTTCGGCTACAGACACCCCTGG 840  
Db 3003 TCCAGTGCTTCAACGGGGGCCAGCAACGACCAACCACTACTTTCGGCTACAGACACCCCTGG 3062  
QY 841 GGGTATTTTGTATTTCAACAGAGATTCACCTGCACACTTTTCCACCAGCTGACTGGCAGGACTC 900  
Db 3063 GGGTATTTTGTATTTCAACAGAGATTCACCTGCACACTTTTCCACCAGCTGACTGGCAGGACTC 3122  
QY 901 ATCAACAACAATTTGGGATTCGGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA 960  
Db 3123 ATCAACAACAATTTGGGATTCGGGCCCAAGAGACTCAACTTTCAAACTCTTCAACATCCAA 3182  
QY 961 GTCAGGAGGTTCACGACGAATGATGGGTTCACACCACTCGCTTAACCTTTACCAAGCAG 1020  
Db 3183 GTCAGGAGGTTCACGACGAATGATGGGTTCACACCACTCGCTTAACCTTTACCAAGCAG 3242  
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACGAGCTTCCGTACGCTCCTCGGCTCTGCGCACCG 1080  
Db 3243 GTTCAAGTCTTCTCGGACTCGGAGTACGAGCTTCCGTACGCTCCTCGGCTCTGCGCACCG 3302  
QY 1081 GGTGCTCCTCTCGGCTTCCGGCGGACGCTGTTTCAATTCGTCGCAATACGGCTACCTGACG 1140  
Db 3303 GGTGCTCCTCTCGGCTTCCGGCGGACGCTGTTTCAATTCGTCGCAATACGGCTACCTGACG 3362  
QY 1141 CTCACAATGGCAGCCCAAGCGTGGAGCTTTCATCTCTTTTACTGCTCGCAATATTTCCCT 1200  
Db 3363 CTCACAATGGCAGCCCAAGCGTGGAGCTTTCATCTCTTTTACTGCTCGCAATATTTCCCT 3422  
QY 1201 TCTCAGATGCTTGAGAAACGGGCAACAACTTTTACCTTTCAGCTACACCTTTTGAGGAAGTGCT 1260  
Db 3423 TCTCAGATGCTTGAGAAACGGGCAACAACTTTTACCTTTCAGCTACACCTTTTGAGGAAGTGCT 3482  
QY 1261 TTCCACAGCAGCTACGCGCAACGAGCAGCTGAGCCGGCTGATGAATCTCTCATTCGAC 1320  
Db 3483 TTCCACAGCAGCTACGCGCAACGAGCAGCTGAGCCGGCTGATGAATCTCTCATTCGAC 3542  
QY 1321 CAATACCTGTATTACTCGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACGAAGAC 1380  
Db 3543 CAATACCTGTATTACTCGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACGAAGAC 3602  
QY 1381 TTGCTGTTTAGCCGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACCT 1440  
Db 3603 TTGCTGTTTAGCCGGTCTCCAGCTGGCATGTCTGTTCAGCCCAAAACCTGGCTACCT 3662  
QY 1441 GGAACCTGTTATTCGCGCAGCAGCGCTTTCTAAAACAAAGAACAGCAACAAAGCAAT 1500  
Db 3663 GGAACCTGTTATTCGCGCAGCAGCGCTTTCTAAAACAAAGAACAGCAACAAAGCAAT 3722  
QY 1501 TTTTACCTCGACTGGTGTCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCT 1560  
Db 3723 TTTTACCTCGACTGGTGTCTTCAAAATATAACCTCAATGGGCGTGAATCCATCATCAACCT 3782  
QY 1561 GGCAGCTGTATGGCCCTCACAAAAGACGACGAAAGCAAGTTCCTTCCCATGAGCGGTGTC 1620  
Db 3783 GGCAGCTGTATGGCCCTCACAAAAGACGACGAAAGCAAGTTCCTTCCCATGAGCGGTGTC 3842  
QY 1621 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGTGCAATTCGATGATT 1680  
Db 3843 ATGATTTTGGAAAAGAGAGCGCGGAGCTTCAAAACATGTGCAATTCGATGATT 3902  
QY 1681 ACAGACGAGAGGAAATTTAAAGCCACTAAACCTGTGGCCACCGAAGATTTGGGACCGTG 1740  
Db 3903 ACAGACGAGAGGAAATTTAAAGCCACTAAACCTGTGGCCACCGAAGATTTGGGACCGTG 3962  
QY 1741 GCAGTCAATTTTCCAGCAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCAATCTATGGGA 1800  
Db 3963 GCAGTCAATTTTCCAGCAGCAGCAGCAGACCCCTGCGACCGGAGATGTGCAATCTATGGGA 4022  
QY 1801 GATTTACCTGGCATGGTGTGGCAAGATAGAGAGCTGTACCTGCAAGGTTCCCATTTGGGCC 1860  
Db 4023 GATTTACCTGGCATGGTGTGGCAAGATAGAGAGCTGTACCTGCAAGGTTCCCATTTGGGCC 4082  
QY 1861 AAAATTTCTTCACACAGATGGACACTTTTCAACCGCTCTCTCTTTATGGGCGGCTTTGAGACTC 1920

Db	4083	AAAAATTCACACAGATGACACTTTTCAACCGCTCTCTCTATATGGCGCGCTTGGACTC	4142
Qy	1921	AAGAAACCGGCTCCTCAGATCCTCATCAAAACACAGCGTGTCTCGAATCCTCGGGC	1980
Db	4143	AAGAACCGGCTCCTCAGATCCTCATCAAAACACAGCGTGTCTCGAATCCTCGGGC	4202
Qy	1981	GAGTTTTTTCAGCTACAAAGTTTGCTTCATTCATCAACCCCAATATCTCCACAGACAAGTGAGT	2040
Db	4203	GAGTTTTTTCAGCTACAAAGTTTGCTTCATTCATCAACCCCAATATCTCCACAGACAAGTGAGT	4262
Qy	2041	GTGGAATTCGATGGGAGCTGCAAGAAAGAAAACAGCAAGCGCTGGAAATCCCGAAGTCGAC	2100
Db	4263	GTGGAATTCGATGGGAGCTGCAAGAAAGAAAACAGCAAGCGCTGGAAATCCCGAAGTCGAC	4322
Qy	2101	TACACATCCAAATATGCAAAATCTGCCAAGTTGATTTTACTGTGGACAAATAGCACTT	2160
Db	4323	TACACATCCAAATATGCAAAATCTGCCAAGTTGATTTTACTGTGGACAAATAGCACTT	4382
Qy	2161	TATACTGAGCTCGCCCCCATTTGGACCCGTTACCTTACCGCTCCCTGT	2209
Db	4383	TATACTGAGCTCGCCCCCATTTGGACCCGTTACCTTACCGCTCCCTGT	4431
RESULT 7			
AX753251			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 2209; Conservative			
0; Mismatches			
0; Indels			
0; Gaps			
0;			
Qy	1	ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC	60
Db	2223	ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC	2282
Qy	61	GAGTGGTGGACATTGAACCTCTGGAGCCCGAAGCCCAAGCCCAACCAAGCAAGCAGGAC	120
Db	2283	GAGTGGTGGACATTGAACCTCTGGAGCCCGAAGCCCAAGCCCAACCAAGCAGGAC	2342
Qy	121	GACGGCCGGGTCTGGTCTTCCTGGCTACAAGTACCTTCGAGACCTTCAACGGACTCGAC	180
Db	2343	GACGGCCGGGTCTGGTCTTCCTGGCTACAAGTACCTTCGAGACCTTCAACGGACTCGAC	2402
Qy	181	AAGGGGAGCCGTCACAGCGGGCGGACGACGCGCCCTCGAGCAGCAAGGCTTACGAC	240
Db	2403	AAGGGGAGCCGTCACAGCGGGCGGACGACGCGCCCTCGAGCAGCAAGGCTTACGAC	2462
Qy	241	CAGCAGCTCAAGCGGGTGACAAATCCGTACTCGGTATTAACACACGCGCGCGAGTTT	300
Db	2463	CAGCAGCTCAAGCGGGTGACAAATCCGTACTCGGTATTAACACACGCGCGCGAGTTT	2522
Qy	301	CAGGAGCGTCTGCAAGAGATACGTCCTTTGGGGGCAACCTTCGGGCGAGCAGTCTCCAG	360

Db	2523	CAGGAGCGTCTGCAAGAGATAGTCTTTTGGGGGCAACCTTCGGGCGAGCAGTCTTCCAG	2582
Qy	361	GCCAAAGACGGGTTTCTCGAACCTCTCGTCTGGTTGAGGAAGCGCTTAAGACGGCTCCT	420
Db	2583	GCCAAAGACGGGTTTCTCGAACCTCTCGTCTGGTTGAGGAAGCGCTTAAGACGGCTCCT	2642
Qy	421	GGAAAGAAACGTCGGTAGAGCAGTCCCAAGAGCAGACTCCCTCGGCGATCGGC	480
Db	2643	GGAAAGAAACGTCGGTAGAGCAGTCCCAAGAGCAGACTCCCTCGGCGATCGGC	2702
Qy	481	AGACAGGCGCAGCAGCCCTCTAAAGAGAGCTCAATTTTGGTCAGACTGGCGACTCAGAG	540
Db	2703	AGACAGGCGCAGCAGCCCTCTAAAGAGAGCTCAATTTTGGTCAGACTGGCGACTCAGAG	2762
Qy	541	TCAGTCCCCGATCCACAACCTCTCGAGAACTCTCCAGAACCCCGCTGCTGGGACCT	600
Db	2763	TCAGTCCCCGATCCACAACCTCTCGAGAACTCTCCAGAACCCCGCTGCTGGGACCT	2822
Qy	601	ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGGAGACAAATACGAAGCGCGGACGGA	660
Db	2823	ACTACAATGGCTTCAGGCGGTGGCGCAACCAATGGGAGACAAATACGAAGCGCGGACGGA	2882
Qy	661	GTGGGTAAATGCCCTCAGGAAATGGCAATTCGATTCACATGGCTGGCGCAGACAGTCAATC	720
Db	2883	GTGGGTAAATGCCCTCAGGAAATGGCAATTCGATTCACATGGCTGGCGCAGACAGTCAATC	2942
Qy	721	ACCACAGCACCCGCACTGGGCTTTGCCCACTCAATAAACAACCTCTCAAGCAAAATC	780
Db	2943	ACCACAGCACCCGCACTGGGCTTTGCCCACTCAATAAACAACCTCTCAAGCAAAATC	3002
Qy	781	TCAGTGTCTTCAACGGGGGCGCAGCAACCACTACTTCGGCTACAGCACCCCTGG	840
Db	3003	TCAGTGTCTTCAACGGGGGCGCAGCAACCACTACTTCGGCTACAGCACCCCTGG	3062
Qy	841	GGGTATTTTGAATTTCAACAGATTCCACTGCCACTTTTCAACAGTCACTGGCAGCGACTC	900
Db	3063	GGGTATTTTGAATTTCAACAGATTCCACTGCCACTTTTCAACAGTCACTGGCAGCGACTC	3122
Qy	901	ATCAACAACAAATTTGGGGATTTCGGGCGCAAGAGACTCAAACTTCAAACTCTTCAACATCAA	960
Db	3123	ATCAACAACAAATTTGGGGATTTCGGGCGCAAGAGACTCAAACTTCAAACTCTTCAACATCAA	3182
Qy	961	GTCAAGAGGTCAAGCAATGATGGCGTCACAACCATCGCTAATACCTTACGAGCAGC	1020
Db	3183	GTCAAGAGGTCAAGCAATGATGGCGTCACAACCATCGCTAATACCTTACGAGCAGC	3242
Qy	1021	GTTCAGAGTCTTCTCGGACTCGGAGTACCACTTCCGTACGTCTCGGCTCTCGCAGCAG	1080
Db	3243	GTTCAGAGTCTTCTCGGACTCGGAGTACCACTTCCGTACGTCTCGGCTCTCGCAGCAG	3302
Qy	1081	GGCTGCTCCTCCGTTCCCGCGGACGTGTTTCAATTCGCAATAACGCTACCTGACG	1140
Db	3303	GGCTGCTCCTCCGTTCCCGCGGACGTGTTTCAATTCGCAATAACGCTACCTGACG	3362
Qy	1141	CTCAAACAATGGCAGCAAGCGGTGGACGTTTCAATTCCTTTTACTGCTGGAAATATTTCCCT	1200
Db	3363	CTCAAACAATGGCAGCAAGCGGTGGACGTTTCAATTCCTTTTACTGCTGGAAATATTTCCCT	3422
Qy	1201	TCTCAGATGCTGAGACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT	1260
Db	3423	TCTCAGATGCTGAGACGGGCAACACTTTTACCTTCAGCTACACCTTTGAGGAAGTGCCT	3482
Qy	1261	TTCCACAGCAGCTACGCGCACAGCGCTCGAGCGGCTGATGAATCTCTCATCGAC	1320
Db	3483	TTCCACAGCAGCTACGCGCACAGCGGCTCGAGCGGCTGATGAATCTCTCATCGAC	3542
Qy	1321	CAATACTGTTATTAAGTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAGGAC	1380
Db	3543	CAATACTGTTATTAAGTGAACAGAACTCAAAATCAGTCCGGAAGTGCCCAAAACAAGGAC	3602
Qy	1381	TTGCTGTTAGCGTGGGTCTCCAGCTGGGATGCTGTTTTCAGGCCCAAAACTGGCTACCT	1440
Db	3603	TTGCTGTTAGCGTGGGTCTCCAGCTGGGATGCTGTTTTCAGGCCCAAAACTGGCTACCT	3662

QY	1441	GGACCTGTTATCGGCAGCGCGTTTCTAAACAAAACAGACAAACACGCAAT	1500		
Db	3663	GGACCTGTTATCGGCAGCGCGTTTCTAAACAAAACAGACAAACACGCAAT	3722		
QY	1501	TTTACCTGGACTGGTGCCTTCAAAATATAACCTCAATGGGGGTGAATCCATCATCAACCT	1560		
Db	3723	TTTACCTGGACTGGTGCCTTCAAAATATAACCTCAATGGGGGTGAATCCATCATCAACCT	3782		
QY	1561	GGCACTGCTATGCGCTCACACAAAGACGACGAAGCAAGTTCTTTCCATGAGCGGTGC	1620		
Db	3783	GGCACTGCTATGCGCTCACACAAAGACGACGAAGCAAGTTCTTTCCATGAGCGGTGC	3842		
QY	1621	ATGATTTTTGGAAAAGAGCGCGGAGCTTCAACACATGCGATGACATGCTGATTT	1680		
Db	3843	ATGATTTTTGGAAAAGAGCGCGGAGCTTCAACACATGCGATGACATGCTGATTT	3902		
QY	1681	ACAGACGAAGAGAAATTAAGCCATTAACCTCTGGCCACCAAGAAATTTGGACCGGTG	1740		
Db	3903	ACAGACGAAGAGAAATTAAGCCATTAACCTCTGGCCACCAAGAAATTTGGACCGGTG	3962		
QY	1741	CGAGTCAATTCAGAGCAGCAGCAGACACACCTTGCACCGGAGATGTGATGGA	1800		
Db	3963	CGAGTCAATTCAGAGCAGCAGCAGACACACCTTGCACCGGAGATGTGATGGA	4022		
QY	1801	GCAATACCTGGCATGCTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTGGGCC	1860		
Db	4023	GCAATACCTGGCATGCTGGCAAGATAGAGACGTGTACCTGACGGGTCCCATTTGGGCC	4082		
QY	1861	AAATTCCTCACACAGATGGACACTTTTCACCCGCTCTCTTATGGCGGCTTTGACTC	1920		
Db	4083	AAATTCCTCACACAGATGGACACTTTTCACCCGCTCTCTTATGGCGGCTTTGACTC	4142		
QY	1921	AGAAACCGGCTCTCAGATCTCATCAAAACACACGCTGTTCTCGAATCTCCGGCG	1980		
Db	4143	AGAAACCGGCTCTCAGATCTCATCAAAACACACGCTGTTCTCGAATCTCCGGCG	4202		
QY	1981	GAGTTTTCAGTCAAGTTTGCTTCAATCATCAACCAATCTCCAGACAGTGAAT	2040		
Db	4203	GAGTTTTCAGTCAAGTTTGCTTCAATCATCAACCAATCTCCAGACAGTGAAT	4262		
QY	2041	GTGGAATTAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCGAAATGCGAG	2100		
Db	4263	GTGGAATTAATGGAGCTGCAGAAAGAAAACAGCAAGCGCTGGAATCCGAAATGCGAG	4322		
QY	2101	TACACATCCAATATGCAAAATCTGCCAAGTTGATTTTACTGTGGACAAATGACTT	2160		
Db	4323	TACACATCCAATATGCAAAATCTGCCAAGTTGATTTTACTGTGGACAAATGACTT	4382		
QY	2161	TATACCTGAGCTCGCCCATTTGGACCGGTACCTTACCGTCCCTGT	2209		
Db	4383	TATACCTGAGCTCGCCCATTTGGACCGGTACCTTACCGTCCCTGT	4431		
RESULT 8					
AF063497					
LOCUS	Adeno-associated virus 1, complete genome.				
DEFINITION	AF063497				
ACCESSION	AF063497.1				
VERSION	GI:4689096				
KEYWORDS	Adeno-associated virus 1				
SOURCE	Adeno-associated virus 1				
ORGANISM	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.				
REFERENCE	1 (bases 1 to 4718)				
AUTHORS	Xiao, W., Chirmule, N., Berta, S.C., McCullough, B., Gao, G. and Wilson, J.M.				
TITLE	Gene therapy vectors based on adeno-associated virus type 1				
JOURNAL	J. Virol. 73 (5), 3994-4003 (1999)				
PUBMED	10196295				
REFERENCE	2 (bases 1 to 4718)				
AUTHORS	Xiao, W. and Wilson, J.M.				
TITLE	Direct Submission				

JOURNAL	Submitted (05-MAY-1998) IHGT, University of Pennsylvania, 3601 Spruce Street, Philadelphia, PA 19104, USA	
FEATURES	Location/Qualifiers	
source	1. .4718	
	/organism="Adeno-associated virus 1"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:85106"	
	/note="AAV1"	
	335. .2206	
CDS	/codon_start=1	
	/product="nonstructural protein"	
	/protein_id="AAD27758.1"	
	/db_xref="GI:4689098"	
	/translation="MPGEYEVIVKPSDLDEHLPGISDSFVSWAEKEWELPPDSMDMLIQEALPTVAELKLDLFLVOMRRVSKAPEALFFVQFKGESYFHLHLVETGVYKLVNLFQSLQIDKLQVTTIRGIEPTLPNFAVTKTRNGAGGKVKVDECIYINYLPGKTOPELQAWNTMEYISACLNLAERKLAHLTVSQTQEQNKSNPNSDAPVIRSKTSAFYMLVGLVDRIKTSQKQEDQASYSFNAASNSRSLKAALDNAGKIMALTSAADYLVGPAPPADIKTNRIYRILELNGVPAVAGSVFLGMAOKRFKRNITWIFGPATTKTINIAENIAHAHVPPYGVVNTNENFPENDCVDRKVIWEEGSKMTAKVYESAKAILGSKRVDDQCKSSAQIDPTPIVITNTNNCAVIDGNTTTFHQOQLQDRMFKFELTRLBHDGKVTQKVEFFRWAQDHTEVAHEFYVRGKANKRAPDADAPKSEPKRACPSADPSTSDAEGAPVDFADRYQNKCSRHAGMLQMLPCKTCERMNQNFICFTHGTRDCSCFPQSVSQPVVRKRTYRKLCAIHLHGRAPEIACSAACDLVNVLDLDCVSEQ"	
	2223. .4433	
CDS	/codon_start=1	
	/product="capsid protein"	
	/protein_id="AAD27757.1"	
	/db_xref="GI:4689097"	
	/translation="MAADGYLPDMLDNLSEIGREBMDLKPAGPKPKANQKQDDRG LVLPGYKVLGPNGLDGEPVNAADAALEHDKAYDQOLKAGDNPYLRYNHADAEFOE RLEQTSFGNGLGRAVFOAKKRVLPLGLVEEGAKTAPGKKRPVQSPQBPSSSGIG KTGQPAKKRLNFGQTSVSDPDPGLGEPATPAAGVPTTMASGGGAPMADNNEGA DTGVNAGNMHCDSTWGLDRVITTRTWALPTNNHLYKQISSASTGVNTDGVFTTIA STPMGYDFNRFHCSPRDMLNNNGFPRKLNFLNIQVKEVNTDGVFTTIA NNLTSYVFSDESYOLPVLSAHQGLCPPPADVFWIPQYGLTYLNNGSAVGRSS FVCLSVFPSQMLRTGNFTFSYTPPEVPHSSVAHSQSLDRIMNPLIDQYLYLARTQ NQSGAQNKDLLFSRGSFAGMSVQPKNMLPGCYRQQRVSRTKTDNNNSNFNTGASK YNLNGRESINFGTAMASHKDDKFFPMGVMIFGKESAGASNTALDNMLITDSEET KATNPVATERFGTAVNFQSSSTDPATGDVHAMGALPGMWQDRDYLQGLPIWAKIPH TDGHPHSPLMGGFLKHPPPQILIKNTVPANPPAEFSATFASFATITQYSGQVSE IEMWLOKENSKRNVPEVQVTSYAKSANVDFVNDNGLYTEPRPIGTRYLTRPL"	
ORIGIN		
Query Match	100.0%; Score 2209; DB 13; Length 4718;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 2209; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
1	ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC	60
2223	ATGGCTGCCGATGGTTATCTTCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTCGC	2282
61	GAGTGTGGGACTTGAAACCTCGAGCCCGAAGCCCAAGCCAAACAGCAAGAGCAGGAC	120
2283	GAGTGTGGGACTTGAAACCTCGAGCCCGAAGCCCAAGCCAAACAGCAAGAGCAGGAC	2342
121	GACGCGCGGGTCTGGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGAGCTCGAC	180
2343	GACGCGCGGGTCTGGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGAGCTCGAC	2402
181	AAGGGGAGCCGCTCAACGCGGGGAGCGAGCGGCCCTCGAGCAGCAGCAGGCTACGAC	240
2403	AAGGGGAGCCGCTCAACGCGGGGAGCGAGCGGCCCTCGAGCAGCAGCAGGCTACGAC	2462
241	CAGCAGCTCAAGCGGGTGACAAATCCGTACCTCGGGTATAACCAACGCCGCGGAGTTT	300
2463	CAGCAGCTCAAGCGGGTGACAAATCCGTACCTCGGGTATAACCAACGCCGCGGAGTTT	2522
301	CAGGAGCGTCTGCAAGAAGATACGCTTTTGGGGGCAACCTCGGGGAGCAGCTCTCCAG	360
2523	CAGGAGCGTCTGCAAGAAGATACGCTTTTGGGGGCAACCTCGGGGAGCAGCTCTCCAG	2582
361	GCCAAGAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGCGCGCTAAGACGGCTCCT	420



Query Match	Best Local Similarity	99.7%; Score 2202.6; DB 6; Length 7447;	Matches 2205; Conservative	0; Mismatches	4; Indels	0; Gaps	0;
Qy	1	ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCAATTGCG	60				
Db	4621	ACGGCTGCCGACGGTTATCTACCCGATTGGCTCGAGGACAACCTCTCTGAGGGCAATTGCG	4680				
Qy	61	GAGTGGTGGGACTTGAAACCTCGAGCCCGGAAGCCCAAGCCAAAGCCAAACGAGCAAAAGCAGGAC	120				
Db	4681	GAGTGGTGGGACTTGAAACCTCGAGCCCGGAAGCCCAAGCCAAACGAGCAAAAGCAGGAC	4740				
Qy	121	GACGGCCGGGGTCTGGTGTCTCTCGCTACAGTACTCGAACCTTCAACCGGACTCGAC	180				
Db	4741	GACGGCCGGGGTCTGGTGTCTCTCGCTACAGTACTCGAACCTTCAACCGGACTCGAC	4800				
Qy	181	AGGGGGAGCCGTCACCGGGGAGCAGCAGGGGCCCTCGAGCAGCAGAGGCCCTAGCAG	240				
Db	4801	AGGGGGAGCCGTCACCGGGGAGCAGCAGGGGCCCTCGAGCAGCAGAGGCCCTAGCAG	4860				
Qy	241	CAGCAGCTCAAGACGGGTGACAAATCCGTACCTCGGTATACCGCCGAGCCGCGAGTTT	300				
Db	4861	CAGCAGCTCAAGACGGGTGACAAATCCGTACCTCGGTATACCGCCGAGCCGCGAGTTT	4920				
Qy	301	CAGGAGCGTCTGCAAGAGATACGCTTTTGGGGCAACCTCGGGCAGCAGTCTTCCAG	360				
Db	4921	CAGGAGCGTCTGCAAGAGATACGCTTTTGGGGCAACCTCGGGCAGCAGTCTTCCAG	4980				
Qy	361	GCCAGAAGCGGGTCTCGAACCTCTCGTCTGGTTGAGGAGGCGCTTAAGACGGCTCCT	420				
Db	4981	GCCAGAAGCGGGTCTCGAACCTCTCGTCTGGTTGAGGAGGCGCTTAAGACGGCTCCT	5040				
Qy	421	GGAAAGAAACGTCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCGGGCATCGGC	480				
Db	5041	GGAAAGAAACGTCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCGGGCATCGGC	5100				
Qy	481	AAGACAGGCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTTCAGACTGGCGACTCAGAG	540				
Db	5101	AAGACAGGCAGCAGCCCGCTAAAAGAGACTCAATTTTGGTTCAGACTGGCGACTCAGAG	5160				
Qy	541	TCAGTCCCAGATCCACAACCTCTCGGAGAACCTTCAGCAACCCCGCTGCTGTGGGACCT	600				
Db	5161	TCAGTCCCAGATCCACAACCTCTCGGAGAACCTTCAGCAACCCCGCTGCTGTGGGACCT	5220				
Qy	601	ACTACAAATGGCTTCAGGCGGTGGCGCCCAATAGGCAGACAATAACGAAGGGCGCCGAG	660				
Db	5221	ACTACAAATGGCTTCAGGCGGTGGCGCCCAATAGGCAGACAATAACGAAGGGCGCCGAG	5280				
Qy	661	GTGGGTAATGCCCTCAGGAAATTTGGCATTTGCGATTCACATGGCTGGGCGACAGAGTCAATC	720				
Db	5281	GTGGGTAATGCCCTCAGGAAATTTGGCATTTGCGATTTCCACATGGCTGGGCGACAGAGTCAATC	5340				
Qy	721	ACACACAGCACCGGACCTGGGGCTTGGCCACTCAATTAACCACTCTACAGCAAAATC	780				
Db	5341	ACACACAGCACCGGACCTGGGGCTTGGCCACTCAATTAACCACTCTACAGCAAAATC	5400				
Qy	781	TCAGTGTCTTCAACGGGGCGCAGCAAGCAACCACTACTTGGCTCAGCACCCCTGG	840				
Db	5401	TCAGTGTCTTCAACGGGGCGCAGCAAGCAACCACTACTTGGCTCAGCACCCCTGG	5460				
Qy	841	GGGTATTTTGTATTCACAGATTCACATGCCACTTTTCAACACGTGACTGGCAGCGACTC	900				
Db	5461	GGGTATTTTGTATTCACAGATTCACATGCCACTTTTCAACACGTGACTGGCAGCGACTC	5520				
Qy	901	ATCAACAAATTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCAA	960				
Db	5521	ATCAACAAATTTGGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCAA	5580				
Qy	961	GTCAAGGGGTACGACGAATGATGGCGTCAACCAATCGCTTAATACTTACCAGGACG	1020				
Db	5581	GTCAAGGGGTACGACGAATGATGGCGTCAACCAATCGCTTAATACTTACCAGGACG	5640				
Qy	1021	GTTCAGAGTCTTCTCGAGCTCGGAGTACCAAGCTTCGGTACGCTCTCGGCTCTCGCACCAG	1080				







QY	1681	ACAGCAGGAGGAAATTAAGCCCACTAACCCCTGTGGCCACCGAAGATTTGGGACCGTG	1740
Db	1681	ACAGCAGGAGGAAATTAAGCCCACTAACCCCTGTGGCCACCGAAGATTTGGGACCGTG	1740
QY	1741	GCAGTCAATTTCCAGAGCAGCAGCAGACCTCGCGACCGAGAGTGTGCATCTATGGGA	1800
Db	1741	GCAGTCAATTTCCAGAGCAGCAGCAGACCTCGCGACCGAGAGTGTGCATCTATGGGA	1800
QY	1801	GCATTAACCTGGATGGTGGCAAGATAGAGCGTGTACCTCGACGGTCCCATTTGGGCC	1860
Db	1801	GCATTAACCTGGATGGTGGCAAGATAGAGCGTGTACCTCGACGGTCCCATTTGGGCC	1860
QY	1861	AAATTCCTCACACAGATGACATTTTACCCGCTCTCTTATGGGCGGCTTTGGACTC	1920
Db	1861	AAATTCCTCACACAGATGACATTTTACCCGCTCTCTTATGGGCGGCTTTGGACTC	1920
QY	1921	AAGAACCCTCCTCAGATCTCATCAAAAACAGCCGTGTTCTCGCAATCCTCGGGG	1980
Db	1921	AAGAACCCTCCTCAGATCTCATCAAAAACAGCCGTGTTCTCGCAATCCTCGGGG	1980
QY	1981	GAGTTTTCAGCTACAAAGTTTCTTCAATTCATCAACCAATCTCCACAGGACAAGTGA	2040
Db	1981	GAGTTTTCAGCTACAAAGTTTCTTCAATTCATCAACCAATCTCCACAGGACAAGTGA	2040
QY	2041	GTGGAAATTAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGATCCCAAGTGCAG	2100
Db	2041	GTGGAAATTAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGATCCCAAGTGCAG	2100
QY	2101	TACACATCAATATGCAAAATCTGCAACAGTTGATTTTACTGTGGACAAATGGACTT	2160
Db	2101	TACACATCAATATGCAAAATCTGCAACAGTTGATTTTACTGTGGACAAATGGACTT	2160
QY	2161	TATACTGAGCCTGGCCCCATGGCACCCGTTACCTTACCGTCCCTCGT	2209
Db	2161	TATACTGAGCCTGGCCCCATGGCACCCGTTACCTTACCGTCCCTCGT	2209
RESULT 11			
AY530611			
LOCUS			
DEFINITION	AY530611 2211 bp DNA linear VRL 24-JUN-2004		
ACCESSION	Adeno-associated virus isolate hu.48 capsid protein Vp1 (cap) gene, complete cds.		
VERSION	AY530611		
KEYWORDS	AY530611.1 GI:46487868		
SOURCE	Adeno-associated virus		
ORGANISM	Adeno-associated virus		
REFERENCE	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.		
AUTHORS	1 (bases 1 to 2211)		
TITLE	Gao,G.; Vandenbergh,L.H., Alvira,M.R., Lu,Y., Calcedo,R., Zhou,X. and Wilson,J.M.		
JOURNAL	Clades of Adeno-associated viruses are widely disseminated in human tissues		
PUBMED	J. Virol. 78 (12), 6381-6388 (2004)		
AUTHORS	15163731		
TITLE	Gao,G.; Vandenbergh,L.H., Alvira,M.R., Lu,Y., Calcedo,R., Zhou,X. and Wilson,J.M.		
JOURNAL	Direct Submission		
FEATURES	Submitted (20-JAN-2004) Gene Therapy Program, Division of Medical Genetics, Department of Medicine, University of Pennsylvania, 3601 Spruce Street, Wistar Institute, Philadelphia, PA 19104, USA		
source	Location/Qualifiers		
	1. .2211		
	/organism="Adeno-associated virus"		
	/mol_type="genomic DNA"		
	/isolate="hu.48"		
	/isolation_source="human tissue"		
	/db_xref="taxon:272636"		
	/country="USA"		
	/notes="clade A"		
gene	1. .2211		

CDS			
	/gene="cap"		
	1. .2211		
	/gene="cap"		
	/codon_start=1		
	/product="capsid protein Vp1"		
	/protein_id="AAS9296.1"		
	/db_xref="GI:46487869"		
	/translation="MAADGYLPDWLEDSLSEIREWDLKPGAPKPKANQOKQDDGRG		
	LVLYKYLGFNGLDGEVPVNAADAALEHDKAYDQQLKAGDNPYLRYNHADAEFQE		
	RLQBDTFGGNLRGAVFOAKRVLLEPLGLVEEGATAPGKKRPVEQSPQBPDSGGIG		
	KTQQPAKRLNFGQGTGDSVDPQLGPPATPAAYPTTMAAGGAPNADNHGGA		
	DVGNAAGNHCDSFSDRWLVTSTRTWALPTYNHLYKQISSTSTGASNDHYFGY		
	GTNGYFDENRFCHFSRDMORLINNNMWRPKRLNFKLFNIQVVEVTTNDVTTIA		
	NNLTSTVQVSDSEYQLPYVLGSAHQGLPPFPADVPMIQQYGLYTLNNGSOAVGRSS		
	PYCLBYFQQMLRTGNFTFSYTFEVPFHSYSQSLSRDKMLPLDQYLYLNRQT		
	NQSGAQNKKLLFSRGPAGMSVQPKMLPCYQQRQSRKMTDNNNSNFTWTSK		
	YLNNGRESINPTGAVANFQSSDTPATDGHAMGALPGVMQMDRVLQGIWAKIPH		
	KATNPVIERFTGTAVANFQSSDTPATDGHAMGALPGVMQMDRVLQGIWAKIPH		
	TDGHFHPSPLMGFGFLANPPQIILIKNTPPVNPAPBFSATKPFASFTQYGTGVSV		
	IEWELQKENSXRNWPEVQYTSNYSANVDTVDNNGLYTEPRIGTRYLRLPL"		
ORIGIN			
	Query Match 99.4%; Score 2196.2; DB 13; Length 2211;		
	Best Local Similarity 99.6%; Pred. No. 0;		
	Matches 2201; Conservative 0; Mismatches 8; Indels 0; Gaps 0;		
QY	1	ATGGCTGCGATGTTATCTTCCAGATTGGCTCGAGGCAACCTCTCTGAGGGCATTCGC	60
Db	1	ATGGCTGCGATGTTATCTTCCAGATTGGCTCGAGGCAACCTCTCTGAGGGCATTCGC	60
QY	61	GAGTGTGCGGACTTGAAACCTGGAGCCCGGAGCCCAAGCCCAACCAAGCAAGCAGGAC	120
Db	61	GAGTGTGCGGACTTGAAACCTGGAGCCCGGAGCCCAAGCCCAACCAAGCAAGCAGGAC	120
QY	121	GACGCCCGGGTCTGGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC	180
Db	121	GACGCCCGGGTCTGGTCTTCTGGCTACAAGTACCTCGGACCTTCAACGGACTCGAC	180
QY	181	AAGGGGAGCCGCTCAACCGCGCGAGCGAGCGGCCCTCGAGCAGCAAGCCCTACGAC	240
Db	181	AAGGGGAGCCGCTCAACCGCGCGAGCGAGCGGCCCTCGAGCAGCAAGCCCTACGAC	240
QY	241	CAGCAGCTCAAGCGGGTGACAATCCGTACCTCGCGTATAACCGCGCAGCGCCAGTTT	300
Db	241	CAGCAGCTCAAGCGGGTGACAATCCGTACCTCGCGTATAACCGCGCAGCGCCAGTTT	300
QY	301	CAGGAGCGTCTGCAAGAAGATACGTCCTTTTGGGGCAACCTCGGCGGAGCAGTCTTCCAG	360
Db	301	CAGGAGCGTCTGCAAGAAGATACGTCCTTTTGGGGCAACCTCGGCGGAGCAGTCTTCCAG	360
QY	361	GCCAAAGAGCGGGTCTCGAACCTCTCGTCTGGTTGAGGAAGCGCTTAAGACGGCTCT	420
Db	361	GCCAAAGAGCGGGTCTCGAACCTCTCGTCTGGTTGAGGAAGCGCTTAAGACGGCTCT	420
QY	421	GGAAAGAAACGTCGCGGTAGCAGTCGCAACAGAGCAGACTCCTCTCGGGCATTCGCG	480
Db	421	GGAAAGAAACGTCGCGGTAGCAGTCGCAACAGAGCAGACTCCTCTCGGGCATTCGCG	480
QY	481	AAGACAGGCGCAGCAGCCGCTAAAAAGAGACTCAATTTTGGCCAGACTGGCGACTCAGAG	540
Db	481	AAGACAGGCGCAGCAGCCGCTAAAAAGAGACTCAATTTTGGCCAGACTGGCGACTCAGAG	540
QY	541	TCAGTCCCGGATCGCAACCTCTCGAGAACCTTCAGCAACCCCGCTGCTGTGGGACT	600
Db	541	TCAGTCCCGGATCGCAACCTCTCGAGAACCTTCAGCAACCCCGCTGCTGTGGGACT	600
QY	601	ACTTACAAATGGCTTCAGGGCGTGGCGACCAATGGCAGCAATAAGAGCGCCGACGGA	660
Db	601	ACTTACAAATGGCTTCAGGGCGTGGCGACCAATGGCAGCAATAAGAGCGCCGACGGA	660
QY	661	GTGGGTAAATGCTCAGGAAATTTGGCATTCGCAATTCACATGCTGGGCGACAGAGTCATC	720
Db	661	GTGGGTAAATGCTCAGGAAATTTGGCATTCGCAATTCACATGCTGGGCGACAGAGTCATC	720

QY 721 ACCACGAGCACCGGACCTGGGCTTGGCCACCTACAATAACCACTCTACAAGCAAATC 780  
DB 721 ACCACGAGCACCGGACCTGGGCTTGGCCACCTACAATAACCACTCTACAAGCAAATC 780  
QY 781 TCAGTGTCTTCAACGGGGGCGAGCAACGACCACTACTTCGGCTACAGCACCCCTGG 840  
DB 781 TCAGTGTCTTCAACGGGGGCGAGCAACGACCACTACTTCGGCTACAGCACCCCTGG 840  
QY 841 GGTATTTTGAATTTCAACAGATTCACCTGCGACTTTTCAACAGCTGACCTGGAGCGACTC 900  
DB 841 GGTATTTTGAATTTCAACAGATTCACCTGCGACTTTTCAACAGCTGACCTGGAGCGACTC 900  
QY 901 ATCAACAACAATTTGGGGATTCGGGCCCCAAGAGACTCAAACTTCAAACTCTTCAACATCCAA 960  
DB 901 ATCAACAACAATTTGGGGATTCGGGCCCCAAGAGACTCAAACTTCAAACTCTTCAACATCCAA 960  
QY 961 GTCAGGAGGTCAAGAGATGATGGGCTCAACCAATGCTGATTAATACCTTACCGACAG 1020  
DB 961 GTCAGGAGGTCAAGAGATGATGGGCTCAACCAATGCTGATTAATACCTTACCGACAG 1020  
QY 1021 GTTCAAGTCTTCCGAGCTCGGAGTACCACTTCCGATGATTCGCTCGCTCGGACACAG 1080  
DB 1021 GTTCAAGTCTTCCGAGCTCGGAGTACCACTTCCGATGATTCGCTCGCTCGGACACAG 1080  
QY 1081 GGTGCTCTCCCTCCGCTTCCGCGGAGCTGTTTCAATTCGCAATACGCTACCTGAGC 1140  
DB 1081 GGTGCTCTCCCTCCGCTTCCGCGGAGCTGTTTCAATTCGCAATACGCTACCTGAGC 1140  
QY 1141 CTCACAATGGGAGCGGAGCGGAGCTGTTTCAATTCGCTGGAATTTTCCCT 1200  
DB 1141 CTCACAATGGGAGCGGAGCGGAGCTGTTTCAATTCGCTGGAATTTTCCCT 1200  
QY 1201 TCTCAGTCTGAGAGCGGACCACTTTTACCTTCACTACAGCTTCACTTGGAGAGTGCCT 1260  
DB 1201 TCTCAGTCTGAGAGCGGACCACTTTTACCTTCACTACAGCTTCACTTGGAGAGTGCCT 1260  
QY 1261 TTCACAGAGCTTACGCGACGAGCGCTGAGCGCTGATGATTCCTCTCATCGAC 1320  
DB 1261 TTCACAGAGCTTACGCGACGAGCGCTGAGCGCTGATGATTCCTCTCATCGAC 1320  
QY 1321 CAATACTGATTTACCTGAGACAGAACTCAAAATCAGTTCGGAAGTCCCAAAACAGGAC 1380  
DB 1321 CAATACTGATTTACCTGAGACAGAACTCAAAATCAGTTCGGAAGTCCCAAAACAGGAC 1380  
QY 1381 TTGCTGTTTACGCTGGCTTCCAGCTGGCACTGTCTTTCAGCCCAAACTGGCTACCT 1440  
DB 1381 TTGCTGTTTACGCTGGCTTCCAGCTGGCACTGTCTTTCAGCCCAAACTGGCTACCT 1440  
QY 1441 GGAACCTGTTTATCGGACGAGCGGCTTCTTAAACCAAAACAGACCAACACAGCAAT 1500  
DB 1441 GGAACCTGTTTATCGGACGAGCGGCTTCTTAAACCAAAACAGACCAACACAGCAAT 1500  
QY 1501 TTTACCTGGACTGGTGTCTTCAAAATATACCTCAATGGGCGTGAATCCATCATCAACCT 1560  
DB 1501 TTTACCTGGACTGGTGTCTTCAAAATATACCTCAATGGGCGTGAATCCATCATCAACCT 1560  
QY 1561 GGCATGCTATGGCTTACAAAGACGAGCAAGTTCCTTCCATGAGCGGTGTC 1620  
DB 1561 GGCATGCTATGGCTTACAAAGACGAGCAAGTTCCTTCCATGAGCGGTGTC 1620  
QY 1621 ATGATTTTGGAAAAGAGCGCGGAGCTTCAAACTGCTATTTGGACAATGTCATGATT 1680  
DB 1621 ATGATTTTGGAAAAGAGCGCGGAGCTTCAAACTGCTATTTGGACAATGTCATGATT 1680  
QY 1681 ACAGACGAAGAGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTGGGACCGTG 1740  
DB 1681 ACAGACGAAGAGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTGGGACCGTG 1740  
QY 1741 GCAGTCAATTTCCAGACGAGCAAGACCTTCCGACCGGAGATGTGATATGGGA 1800  
DB 1741 GCAGTCAATTTCCAGACGAGCAAGACCTTCCGACCGGAGATGTGATATGGGA 1800

QY 1801 GCATTACCTGGCATGCTGGCAAGATAGAGACGTGTACTCGACGGTCCCATTTGGGCT 1860  
DB 1801 GCATTACCTGGCATGCTGGCAAGATAGAGACGTGTACTCGACGGTCCCATTTGGGCT 1860  
QY 1861 AAAATTCCTCACACAGATGGACATTTTCAACCGCTTCTCTTATGGGCGGCTTTGGACTC 1920  
DB 1861 AAAATTCCTCACACAGATGGACATTTTCAACCGCTTCTCTTATGGGCGGCTTTGGACTC 1920  
QY 1921 AAGAACCCTGCTTCTCAGATCCTCATAAAAACAGCCCTGTTCTTCCGAAATCCTCGGCG 1980  
DB 1921 AAGAACCCTGCTTCTCAGATCCTCATAAAAACAGCCCTGTTCTTCCGAAATCCTCGGCG 1980  
QY 1981 GAGTTTTCAGCTACAAAGTTTGTCTTCAATTCATCAACCAATCTCCACAGGACAAGTGA 2040  
DB 1981 GAGTTTTCAGCTACAAAGTTTGTCTTCAATTCATCAACCAATCTCCACAGGACAAGTGA 2040  
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCAAGTGCAG 2100  
DB 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGACGCTGGAATCCCAAGTGCAG 2100  
QY 2101 TACACATCCAAATATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGACTT 2160  
DB 2101 TACACATCCAAATATATGCAAAATCTGCCAACGTTGATTTTACTGTGGACAACAATGACTT 2160  
QY 2161 TATACCTGAGCTCGCCCAATTCGACCCCTTACCTTACCCGCTCCCTGT 2209  
DB 2161 TATACCTGAGCTCGCCCAATTCGACCCCTTACCTTACCCGCTCCCTGT 2209

RESULT 12  
CS073492  
LOCUS 2214 bp DNA linear PAT 05-MAY-2005  
DEFINITION Sequence 80 from Patent WO2005033321.  
ACCESSION CS073492  
VERSION CS073492.1 GI:63090471  
KEYWORDS .  
SOURCE unidentified  
ORGANISM unclassified.  
REFERENCE 1  
AUTHORS Wilson,J.M., Gao G., Alvira,M.R. and Vandenberghe,L.H.  
TITLE Adeno-associated virus (aav) clades, sequences, vectors containing same, and uses therefor  
JOURNAL Patent: WO 2005033321-A 80 14-APR-2005;  
The Trustees of the University of Pennsylvania (US)  
FEATURES  
source  
1..2214  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3264"  
/note="new AAV serotype, clone hu.43"

ORIGIN

Query Match 96.2%; Score 2124; DB 6; Length 2214;  
Best Local Similarity 97.8%; Pred. No. 0;  
Matches 2164; Conservative 0; Mismatches 45; Indels 3; Gaps 1;

QY 1 ATGGCTGCGCATGGTTATCTTCCAGATTGGCTCGAGCAACCTCTCTGAGGSCATTTCG 60  
DB 1 ATGGCTGCGCATGGTTATCTTCCAGATTGGCTCGAGCAACCTCTCTGAGGSCATTTCG 60  
QY 61 GAGTGTGGGACTTGAACCTTGGAGCCCGCAAGCCCAAGCCCAAGCCCAAGCAGGAG 120  
DB 61 GAGTGTGGGACTTGAACCTTGGAGCCCGCAAGCCCAAGCCCAAGCAGGAG 120  
QY 121 GACGGCGGGGTCTGGTGTCTTCTTGGCTACAGTACTCGACCCCTTCAACGAGCTGAC 180  
DB 121 GACGGCGGGGTCTGGTGTCTTCTTGGCTACAGTACTCGACCCCTTCAACGAGCTGAC 180  
QY 181 AAGGGGAGCCGCTCAACGGCGGACGAGCGGCGCTTCGAGCACCAAGGCTTACGAC 240  
DB 181 AAGGGGAGCCGCTCAACGGCGGACGAGCGGCGCTTCGAGCACCAAGGCTTACGAC 240

Qy	241	CAGCAGCTCAAAGCGGTGACAAATCCGTACTTCGGGTATACCAACGCGCGACGCGGAGTTT	300
Db	241	CAGCAGCTCAAAGCGGTGACAAATCCGTACTTCGGGTATACCAACGCGCGACGCGGAGTTT	300
Qy	301	CAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGGAACTCTGGGCGAGCAGTCTTCCAG	360
Db	301	CAGGAGCGTCTGCAAGAAGATACGCCCTTTTGGGGGGCAACCTCGGGCGAGCAGTCTTCCAG	360
Qy	361	GCCAAAGAAGCGGGTCTTCGAACCTCTCGGTCTGGTTCGAGGAAGCGCTTAAGACGGCTCCT	420
Db	361	GCCAAAGAAGCGGGTCTTCGAACCTCTCGGTCTGGTTCGAGGAAGCTGCTTAAGAAGCTCCT	420
Qy	421	GGAAAGAAACGTCCGGTAGACAGTGCCCACA---AGAGCCAGACTCCTCTCGGGCAGTC	477
Db	421	GGAAAGAAAGAGACCGGTAGAACCGTCACTCAGCGTTCCCGCGACTCCTCCACGGGCAATC	480
Qy	478	GGCAAGACAGCCAGCAGACCCGCTAAAGAGACTCAATTTTGGTTCAGACTGGGGACTCA	537
Db	481	GGCAAGAAAGCCAGCAGACCCGCTAAAGAGAGACTGAACCTTTGGTTCAGACTGGGCACTCA	540
Qy	538	GAGTCAGTCCCGGATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGGA	597
Db	541	GAGTCAGTCCCGGACCTTCAACCAATCGGAGAACCAACAGCAGGCCCTCTGTGTCGGGA	600
Qy	598	CCTACTACAAATGGCTTCAGGCGGTGGCGCACCAATGGCAGACAATAACGAAGCGCGCAC	657
Db	601	TCTGGTACAATGGCTGCAAGCGGTGGCGCTCCAAATGGCAGACAATAACGAAGCGCGCAC	660
Qy	658	GGAGTGGGTAAATGCCTCAGGAAATGGGATTTGGGATTTCCATGGCTGGGGCGACAGATC	717
Db	661	GGAGTGGGTAAATGGCTCAGGAAATTTGGCATTTGGCATTTCCATGGCTGGGCGACAGAGTC	720
Qy	718	ATCACCAACGACACCCGACCTGGCGCTTGCCCAACCTACAATAACCACTCTCTACAAGCAA	777
Db	721	ATCACCAACGACACCCGACCTGGCGCTTGCCCACTTACAATAACCACTCTCTACAAGCAA	780
Qy	778	ATCTCCAGTGTCTTCAACGGGGGCGCAGAAAGACAACCACTACTTCGGCTACAGCACCCCC	837
Db	781	ATCTCCAGTGTCTTCAACGGGGGCGCAGAAAGACAACCACTACTTCGGCTACAGCACCCCC	840
Qy	838	TGGGGGTATTTTGATTTTCAACAGATTCGACCTGGCCATCTTTTCAACGCTGACGTGACGGGA	897
Db	841	TGGGGGTATTTTGATTTTCAACAGATTCGACCTGGCCATCTTTTCAACGCTGACGTGACGGGA	900
Qy	898	CTCATCAACAACAATTCGGGATTCGGGCGCAAGAGACTCAACTTCAACTCTTCAACATC	957
Db	901	CTCATCAACAACAATTCGGGATTCGGGCGCAAGAGACTCAACTTCAAACTCTTCAACATC	960
Qy	958	CAAGTCAAGGAGGTACAGCAAGAAATGATGGGTGTCACCAACATCGCTAATAACCTTACCAAGC	1017
Db	961	CAAGTCAAGGAGGTACAGCAAGAAATGATGGGTGTCACCAACATCGCTAATAACCTTACCAAGC	1020
Qy	1018	ACGTTTCAAGTCTTCTCGGACTCGGAGTACAGCTTCGGTACGTTCCTCGGCTCTGGGCAC	1077
Db	1021	ACGTTTCAAGTCTTCTCGGACTCGGAGTACCAAGTTCGGTACGTTCCTCGGCTCTGGGCAC	1080
Qy	1078	CAGGCTGCCTCCCTCGGTTCCCGCGGAGGTTCATGATTCGCGCAATACGGCTACCTG	1137
Db	1081	CAGGCTGCCTCCCTCGGTTCCCGCGGAGGTTCATGATTCGCGCAATACGGCTACCTG	1140
Qy	1138	ACGCTCAACAATGGCAGCCAGCCGTGGGAGTTCATCTTTTACTGCCTGGAAATATTC	1197
Db	1141	ACGCTCAACAATGGCAGCCAGCCGTGGGAGTTCATCTTTTACTGCCTGGAAATATTC	1200
Qy	1198	CCTTCTCAGATGCTTGAGAAAGGGCAACAACTTTACCTTCAGTCTACACTTTTGAGGAAGTG	1257
Db	1201	CCTTCTCAGATGCTTGAGAAAGGGCAACAACTTTACCTTCAGTCTACACTTTTGAGGAAGTG	1260
Qy	1258	CCTTTCACAGCAGCTACGGCCAGCCAGAGCTCGACCGGCTGATGATCTCTCTCATC	1317
Db	1261	CCTTTCACAGCAGCTACGGCCAGCCAGAGCTCGACCGGCTGATGATCTCTCTCATC	1320
Qy	1318	GACCAATACCTGTATTACCTGAAACAGAACTCAAAATCAGTTCGGGAAGTGCCCAAAACAAG	1377

Db	1321	GTCCAATACCTGTATTACCTGAACAGAACTCAAAATCAGTCGGGAAGTCCCAAAACAAG	1380
Qy	1378	GACCTGCTGTTTAGCGGTCTCAGCTGGCATCTCTGTTCAGGCCCAAAAACACTGGCTA	1437
Db	1381	GACTTGCTGTTCAGCGTGGGTCTCAGCTGGCATGTCTGTTTCAGGCCCAAAAACACTGGCTA	1440
Qy	1438	CCTGGACCCCTGTTATCGGCAGCAGCGGTTCTTAAACCAAAAACAGACAACAACAGC	1497
Db	1441	CCTGGACCCCTGTTATCGGCAGCAGCGGTTCTTAAACCAAAAACAGACAACAACAGC	1500
Qy	1498	AATTTTACCTGGACTGGTGTCTCAAAATATAACCTCAATGGGGGTGAATCCATCATCAAC	1557
Db	1501	AATTTTACCTGGACTGGTGTCTCAAAATATAACCTCAATGGGGGTGAATCCATCATCAAC	1560
Qy	1558	CCTGGCACTGCTATGCTCACACAAGACGACGAAGACAAGTTCTTCCCATGAGCGGT	1617
Db	1561	CCTGGCACTGCTATGCTCACACAAGACGACGAAGACAAGTTCTTCCCATGAGCGGT	1620
Qy	1618	GTCAATGATTTTTGGAAAAAGAGCGCGGAGCTTCAAAACACTGCATTTGGACAATGTCATG	1677
Db	1621	GTCAATGATTTTTGGAAAAAGAGCGCGGAGCTTCAAAACACTGCATTTGGACAATGTCATG	1680
Qy	1678	ATTACAGACGAAGAGAAATTTAAAGCCAATAACCTGTGGGCCACCGAAAGATTTGGGACC	1737
Db	1681	ATTACAGACGAAGAGAAATTTAAAGCCAATAACCTGTGGGCCACCGAAAGATTTGGGACC	1740
Qy	1738	GTGGCAGTCAATTTTCAGAGCAGCAGACAGACCCCTGCGACCGGAGATGTGCATGCTATG	1797
Db	1741	GTGGCAGTCAATTTTCAGAGCAGCAGACAAGACCCCTGCGACCGGAGATGTGCATGCTATG	1800
Qy	1798	GGAGCATTAACCTGGCATTGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCATTGG	1857
Db	1801	GGAGCATTAACCTGGCATTGGTGTGGCAAGATAGAGACGTGTACCTGCAGGGTCCCATTGG	1860
Qy	1858	GCCAAAATTCCTCACACAGATGAGACACTTTTCACCCGTCTCCTCTTATGGCGGCTTGG	1917
Db	1861	GCCAAAATTCCTCACACAGATGAGACACTTTTCACCCGTCTCCTCTTATGGCGGCTTGG	1920
Qy	1918	CTCAAGAACCCGCTCCTCAGATCCTCATCAAAAAACAGCGCTGTCTCTCGGAATCCTCCG	1977
Db	1921	CTCAAGAACCCGCTCCTCAGATCCTCATCAAAAAACAGCGCTGTCTCTCGGAATCCTCCG	1980
Qy	1978	GCAGGTTTTACGCTACAAAGTTTGTCTATTATCAACCCAAATCTCCACAGACAAGTG	2037
Db	1981	GCAGGTTTTACGCTACAAAGTTTGTCTATTATCAACCCAAATCTCCACAGACAAGTG	2040
Qy	2038	AGTGTGGAAATTCATGGGAGCTGCAGAGAGAAAACAGAAAGCGCTGGAATCCGAGTG	2097
Db	2041	AGTGTGGAAATTCATGGGAGCTGCAGAGAGAAAACAGAAAGCGCTGGAATCCGAGTG	2100
Qy	2098	CAGTACACATCCAAATTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGGA	2157
Db	2101	CAGTACACATCCAAATTATGCAAAATCTGCCAAACGTTGATTTTACTGTGGACAAACAATGGA	2160
Qy	2158	CTTTTACTAGCGCTCGCCCATTTGGCACCCGTTTACCTTACCGTCCCTGT	2209
Db	2161	CTTTTACTAGCGCTCGCCCATTTGGCACCCGTTTACCTTACCGTCCCTGT	2212

RESULT 13	AY530606	AY530606	2214 bp	DNA	linear	VRL 24-JUN-2004
LOCUS						
DEFINITION		Adeno-associated virus isolate hu.43 capsid protein VP1 (cap) gene, complete cds.				
ACCESSION		AY530606				
VERSION		AY530606.1	GI:46487858			
KEYWORDS						
SOURCE		Adeno-associated virus				
ORGANISM		Adeno-associated virus				
REFERENCE		Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus. 1 (bases 1 to 2214)				
AUTHORS		Gao, G., Vandenberghe, L.H., Alvirza, M.R., Lu, Y., Calcedo, R., Zhou, X.				



Db 1441 CTTGACCCCTGTTATCGGCAGCAGCGCGTTCTTAAACAAAAACAGACAACAACACAGC 1500

Qy 1498 AATTTTTACCTGACTGGTGGCTTCAAAATATAAATCAATCTCAATGGGGGGAATCCATCATCAAC 1557  
|||||

Db 1501 AATTTTTACCTGACTGGTGGCTTCAAAATATAAATCAATCTCAATGGGGGGAATCCATCATCAAC 1560  
|||||

Qy 1558 CTTGGCAGCTGCTATGGCCTCACACAAAGACGACGAAGCAAGTCTTTTCCCATGAGCGGT 1617  
|||||

Db 1561 CTTGGCAGCTGCTATGGCCTCACACAAAGACGACGAAGCAAGTCTTTTCCCATGAGCGGT 1620  
|||||

Qy 1618 GTCATGATTTTGGAAAAAGAGCGCGGAGCTTCAAAACATCTGCAATTTGACAAATGTCATG 1677  
|||||

Db 1621 GTCATGATTTTGGAAAAAGAGCGCGGAGCTTCAAAACATCTGCAATTTGACAAATGTCATG 1680  
|||||

Qy 1678 ATTAACAGACGAAGAGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTTGGACC 1737  
|||||

Db 1681 ATTAACAGACGAAGAGAAATTAAGCCACTAACCCCTGTGGCCACCGAAAGATTTGGACC 1740  
|||||

Qy 1738 GTGGCAGTCAATTTCCAGAGCAGCAGCAGACAGCCCTGCGACCGGAGATGTCATGCTATG 1797  
|||||

Db 1741 GTGGCAGTCAATTTCCAGAGCAGCAGCAGACAGCCCTGCGACCGGAGATGTCATGCTATG 1800  
|||||

Qy 1798 GGAGCAATTACCTGGCATGGTGTGGCAAGATAGAGAGTGTACTGACAGGCTCCCATTTGG 1857  
|||||

Db 1801 GGAGCAATTACCTGGCATGGTGTGGCAAGATAGAGAGTGTACTGACAGGCTCCCATTTGG 1860  
|||||

Qy 1858 GCCAAATTTCTCACACAGATGGACACTTTTCAACCCGCTCTCTCTTATGGGGGGCTTTGGA 1917  
|||||

Db 1861 GCCAAATTTCTCACACAGATGGACACTTTTCAACCCGCTCTCTCTTATGGGGGGCTTTGGA 1920  
|||||

Qy 1918 CTCAAGAACCCGCTCTCAGATCCCTCATCAAAACACAGCCCTGTTCTCGGAATCCTCG 1977  
|||||

Db 1921 CTCAAGAACCCGCTCTCAGATCCCTCATCAAAACACAGCCCTGTTCTCGGAATCCTCG 1980  
|||||

Qy 1978 GCGAGTTTTTCAGTACAAAGTTTGTCTTATTCATCACCCCAATCTCCACAGACAAAGTG 2037  
|||||

Db 1981 GCGAGTTTTTCAGTACAAAGTTTGTCTTATTCATCACCCCAATCTCCACAGACAAAGTG 2040  
|||||

Qy 2038 AGTGTGGAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGGAATCCGAAAGTG 2097  
|||||

Db 2041 AGTGTGGAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGGAATCCGAAAGTG 2100  
|||||

Qy 2098 CAGTACACATCCAAATTTGCAAAATCTGCAAGTTGATTTTACTGTGCAACAATGGA 2157  
|||||

Db 2101 CAGTACACATCCAAATTTGCAAAATCTGCAAGTTGATTTTACTGTGCAACAATGGA 2160  
|||||

Qy 2158 CTTTATACTGAGCTCGCCCCCATTTGGACCCGTTACTTTACCGTCCCTGT 2209  
|||||

Db 2161 CTTTATACTGAGCTCGCCCCCATTTGGACCCGTTACTTTACCGTCCCTGT 2212  
|||||

## RESULT 14

CS073493  
LOCUS 2211 bp DNA linear PAT 05-MAY-2005  
DEFINITION Sequence 81 from Patent WO2005033321.  
ACCESSION CS073493  
VERSION CS073493.1 GI:63090472

KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
unclassified.

## REFERENCE

1  
AUTHORS Wilson, J.M., Gao, G., Alvira, M.R. and Vandenberghe, L.H.  
TITLE Adeno-associated virus (aav) clades, sequences, vectors containing  
same, and uses therefor  
JOURNAL Patent: WO 2005033321-A 81 14-APR-2005;  
The Trustees of the University of Pennsylvania (US)

## FEATURES

source  
1..2211  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="new AAV serotype, clone hu.44"

## ORIGIN

Query Match 95.5%; Score 2109.8; DB 6; Length 2211;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 2147; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy 1 ATGGCTGCGGATGGTATCTTCCAGATTGGCTCGAGGACACCTCTCTGAGGGCATTTGCG 60  
Db 1 ATGGCTGCGGATGGTATCTTCCAGATTGGCTCGAGGACACTCTCTCTGAAAGGAATAAGA 60  
|||||

Qy 61 GAGTGGTGGGACTTTGAAACCTGGAGCCCGGAAAGCCAAAGCCAAACAGCAAAAGCAGGAC 120  
Db 61 CAGTGGTGGAGCTCAGACCTTGGCCCAACCAACCAAGCCCGAGAGGGCATTAAGGAC 120  
|||||

Qy 121 GACGCCCGGGGTCTGGTGTCTTCTGGCTTACAAGTACCTTCGGAACCTTCAACGGACTCGAC 180  
Db 121 GACAGCAGGGGTCTTGTGCTTCTGGGTACAAGTACCTCGGACCTTCAACGGACTCGAC 180  
|||||

Qy 181 AAGGGGAGCCGTCACACGCGCGGACGACGCGGCCCTCGAGCAGCAAGGCCCTACGAC 240  
Db 181 AAGGGAGAGCCGTCACAGGCGACGCGCGGCCCTCGAGCAGCAAGGCCCTACGAC 240  
|||||

Qy 241 CAGCAGCTCAAAAGCGGGTGACAATCCGTACCTCGGGTATAACCAAGCCGAGCTTT 300  
Db 241 CCGCAGCTCGACGCGGAGACCAACCCGTACCTCAAGTACAACCAAGCCGAGCTTT 300  
|||||

Qy 301 CAGGAGCGTCTGCAAGAAAGATACGCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360  
Db 301 CAGGAGCGCTTAAAGAAAGATACGCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360  
|||||

Qy 361 GCCAAGAGCGGGTCTCGAAGCCTCTCGGTCTGGTTGAGGAGGCGCTAAGCGGCTCCT 420  
Db 361 GCCAAGAGCGGGTCTCGAAGCCTCTCGGTCTGGTTGAGGAGGCGCTGAGAGCGGCTCCT 420  
|||||

Qy 421 GGAAAGAAACGTCCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCTCGGGCATCGGC 480  
Db 421 GGAAAGAAACGTCCGGTAGAGCAGTCGCCACAAGAGCCAGACTCTCTCTCGGGCATCGGC 480  
|||||

Qy 481 AAGAAGCGCCAGCAGCCCGCTTAAAGAGATCTAAATTTTGGTCAAGCTGGCGACTCAGAG 540  
Db 481 AAGAAGCGCCAGCAGCCCGCTTAAAGAGATCTAAATTTTGGTCAAGCTGGCGACTCAGAG 540  
|||||

Qy 541 TCAGTCCCGGATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGGACCT 600  
Db 541 TCAGTCCCGGATCCACAACCTCTCGGAGAACCTCCAGCAACCCCGCTGCTGTGGGACCT 600  
|||||

Qy 601 ACTACATGGCTTCAGGGCGTGGCGCACCAATGGCAGACAATAAGAGGGCGCCAGCGA 660  
Db 601 ACTACATGGCTTCAGGGCGTGGCGCACCAATGGCAGACAATAAGAGGGCGCCAGCGA 660  
|||||

Qy 661 GTGGTAATGCTCAGGAAATTTGGCATTTGGCATTTCCATTCGATTCGATTCGATTCATC 720  
Db 661 GTGGTAATGCTCAGGAAATTTGGCATTTGGCATTTCCATTCGATTCGATTCGATTCATC 720  
|||||

Qy 721 ACCCAGCAGCCCGCAGCTGGGCTTGGCCACCTTACATTAACCACTCTTACAAGCAATC 780  
Db 721 ACCCAGCAGCCCGCAGCTGGGCTTGGCCACCTTACATTAACCACTCTTACAAGCAATC 780  
|||||

Qy 781 TCCAGTGTCTCAACGGGGCCAGCAACGACAACTACTTCGGCTACAGACCCCTCGG 840  
Db 781 TCCAGTGTCTCAACGGGGCCAGCAACGACAACTACTTCGGCTACAGACCCCTCGG 840  
|||||

Qy 841 GGGTATTTTGAATTTCAACAGATTTCACTGGCACTTTTCAACGCTGAGCTGGCAGGACTC 900  
Db 841 GGGTATTTTGAATTTCAACAGATTTCACTGGCACTTTTCAACGCTGAGCTGGCAGGACTC 900  
|||||

Qy 901 ATCAACAACAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCA 960  
Db 901 ATCAACAACAATTTGGGATTTCCGGCCCAAGAGACTCAACTTCAAACTCTTCAACATCCA 960  
|||||

Qy 961 GTCAAGGAGGTCAACGACGAATGATGGCGTCACAACTCGCTTAATACCTTACGAGCAG 1020  
Db 961 GTCAAGGAGGTCAACGACGAATGATGGCGTCACAACTCGCTTAATACCTTACGAGCAG 1020  
|||||







QY 61 GAGTGGTGGGACTTGAACCTTGGAGCCCGAAGCCCAAGCCAAACGAGCAAAAGCAGGAC 120  
DB 61 CAGTGGTGAAGCTCAGACTCTGGCCCAACCAACCAAGCCCGGAGAGCGGATAAGGAC 120  
QY 121 GACGGCCGGGCTCTGGTGTCTCTCGCTACAAAGTACTCTGGACCCCTTCAACGGACTCGAC 180  
DB 121 GACGCGAGGGGTCTTGGTGTCTCTGGGTACAAGTACTCTGGACCCCTTCAACGGACTCGAC 180  
QY 181 AAGGGGAGCCCGTCAACCGCGCGGACGCGAGCGGCCCTCGAGCAGCAAGGCTTACGAC 240  
DB 181 AAGGGAGAGCCCGTCAACGAGGAGAGCGCGCGGCCCTCGAGCAGCAAGGCTTACGAC 240  
QY 241 CAGGAGCTCAAGAGCGGTGACAACTCGTACTCTGGGTATAACCAAGCGCGAGCGGATTT 300  
DB 241 CAGGAGCTCGACAGCGGAGACAAACCGTACTCTCAAGTACAACCAAGCGCGAGCGGATTT 300  
QY 301 CAGGAGCGTCTGCAAGAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360  
DB 301 CAGGAGCGCTTAAAGAGATACGTCTTTTGGGGGCAACCTCGGGCGAGCAGTCTTCCAG 360  
QY 361 GCCAAGAGCGGTTCTCGAACTCTCGTCTGGTTGAGGAAGGCGCTAAGACGGCTCT 420  
DB 361 GCCAAGAGCGGTTCTCGAACTCTCGTCTGGTTGAGGAAGGCGCTGAGACGGCTCT 420  
QY 421 GGAAAGAAACGTCCGGTAGAGCAGTCCGCAAGGGCCAGACTCTCTCTCGGGCATCGGC 480  
DB 421 GGAAAGAAACGTCCGGTAGAGCAGTCCGCAAGGGCCAGACTCTCTCTCGGGCATCGGC 480  
QY 481 AAGACAGGCGCAGCAGCGCCGCTAAAAGAGACTCAATTTTGGTCAAGTGGCGACTCAGAG 540  
DB 481 AAGACAGGCGCAGCAGCGCCGCTAAAAGAGACTCAATTTTGGTCAAGTGGCGACTCAGAG 540  
QY 541 TCAGTCCCAGATCCAACTCTCGGAGAACTTCCAGCAACCCCGCTCTGTGGGACCT 600  
DB 541 TCAGTCCCAGATCCAACTCTCGGAGAACTTCCAGCAACCCCGCTCTGTGGGACCT 600  
QY 601 ACTACAAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATAACGAAGCGCGCAGGA 660  
DB 601 ACTACAAATGGCTTCAGGCGGTGGCGCAACCAATGGCAGACAATAACGAAGCGCGCAGGA 660  
QY 661 GTGGGTAAATGCTTCAGAAATTTGGCATTCGATTCACATGGCTGGGCGACAGAGTCATC 720  
DB 661 GTGGGTAAATGCTTCAGAAATTTGGCATTCGATTCACATGGCTGGGCGACAGAGTCATC 720  
QY 721 ACCACGAGCACCGGCACTTGGGCTTGGCCACCTACAAATAACCACTCTACAAGCAAAATC 780  
DB 721 ACCACGAGCACCGGCACTTGGGCTTGGCCACCTACAAATAACCACTCTACAAGCAAAATC 780  
QY 781 TCCAGTGTCTTCAACGGGGGCCAGCAACGACACCACTACTTCGGCTACAGCAACCCCTGG 840  
DB 781 TCCAGTGTCTTCAACGGGGGCCAGCAACGACACCACTACTTCGGCTACAGCAACCCCTGG 840  
QY 841 GGGTATTTTGAATTCACAGATTCACACTGCACTTTTACACAGTCACTGGCAGGACATC 900  
DB 841 GGGTATTTTGAATTCACAGATTCACACTGCACTTTTACACAGTCACTGGCAGGACATC 900  
QY 901 ATCAACAAATTTGGGGAATTCGGGCCCAAGAGACTCAACTTCAAACTTCTTCAACATCCAA 960  
DB 901 ATCAACAAATTTGGGGAATTCGGGCCCAAGAGACTCAACTTCAAACTTCTTCAACATCCAA 960  
QY 961 GTCAAGGAGGTACGACGAATGATGGGTCAACCACTCGCTAATTAACCTTACCAGCAG 1020  
DB 961 GTCAAGGAGGTACGACGAATGATGGGTCAACCACTCGCTAATTAACCTTACCAGCAG 1020  
QY 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCACTTCCGTGATCGTCTCGGCTCTGGCAGCAG 1080  
DB 1021 GTTCAAGTCTTCTCGGACTCGGAGTACCACTTCCGTGATCGTCTCGGCTCTGGCAGCAG 1080  
QY 1081 GGTGCTCTCCCTCCGTTCGGCGGAGCGTGTTCATGATTCGGAATCCGCAATACGGCTACCTGAC 1140  
DB 1081 GGTGCTCTCCCTCCGTTCGGCGGAGCGTGTTCATGATTCGGAATTCGGAATACGGCTACCTGAC 1140

QY 1141 CTCACAAATGCGACGCGTGGAGCGTTTACTCTTCTTACTGCTCGGAATATTTCCCT 1200  
DB 1141 CTCACAAATGCGAGCGCAAGCCGTGGAGCGTTTACTCTTCTTACTGCTCGGAATATTTCCCT 1200  
QY 1201 TCTCAGATGCTGAGAAACGGGCAACAACTTTTACCTTTCAGCTACACCTTTTGAGGAAGTGCCT 1260  
DB 1201 TCTCAGATGCTGAGAAACGGGCAACAACTTTTACCTTTCAGCTACACCTTTGAGGAAGTGCCT 1260  
QY 1261 TTCCAAGAGCTACGCGCACAGCAGCGCTTGAACCGGCTGATGAATCTCTCATCGAC 1320  
DB 1261 TTCCAAGAGCTACGCGCACAGCAGCGCTTGAACCGGCTGATGAATCTCTCATCGAC 1320  
QY 1321 CAATACCTGTATTTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAGGAC 1380  
DB 1321 CAATACCTGTATTTACCTGAAACAGAACTCAAAATCAGTCCGGAAGTGCCTCAAAACAGGAC 1380  
QY 1381 TTGCTGTTTACGCTGGGTCTCCAGCTGGCATGTCTTTCAGCCCAAACTGCGTACCT 1440  
DB 1381 TTGCTGTTTACGCTGGGTCTCCAGCTGGCATGTCTTTCAGCCCAAACTGCGTACCT 1440  
QY 1441 GGAACCTGTTATCGGCGAGCGGCTTTCTAAAACAAACAAAGAGAGAGCAACAAAGCAAT 1500  
DB 1441 GGAACCTGTTATCGGCGAGCGGCTTTCTAAAACAAACAAAGAGAGAGCAACAAAGCAAT 1500  
QY 1501 TTTTACCTGAGCTGGTCTTCAAAATATAAATTAACCTCAATGGGCGTGAATCCATCATCAACCT 1560  
DB 1501 TTTTACCTGAGCTGGTCTTCAAAATATAAATTAACCTCAATGGGCGTGAATCCATCATCAACCT 1560  
QY 1561 GGCACCTGTCTGCGCTCACACAAAGACGACGAAAGTCTTTCCTTCCATGAGCGGTGTC 1620  
DB 1561 GGCACCTGTCTGCGCTCACACAAAGACGACGAAAGTCTTTCCTTCCATGAGCGGTGTC 1620  
QY 1621 ATGATTTTGGAAAGAGAGCGCGGAGCTTCAAACACTGTGATTTGGGACAAATGTCTATGAT 1680  
DB 1621 ATGATTTTGGAAAGAGAGCGCGGAGCTTCAAACACTGTGATTTGGGACAAATGTCTATGAT 1680  
QY 1681 ACAGACGAGAGGAAATTAAGCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTG 1740  
DB 1681 ACAGACGAGAGGAAATTAAGCCACTAACCTGTGGCCACCGAAGATTTGGGACCGTG 1740  
QY 1741 GCAGTCAATTTTCAGAGCAGCAGCAACAGCTCTGCGACCGGAGATGTGATATGGGA 1800  
DB 1741 GCAGTCAATTTTCAGAGCAGCAGCAACAGCTCTGCGACCGGAGATGTGATATGGGA 1800  
QY 1801 GCAATACCTGGCATGGTGTGGCAAGGTAGAGACGTGTACCTGAGGGTCCCAATTTGGGCG 1860  
DB 1801 GCAATACCTGGCATGGTGTGGCAAGGTAGAGACGTGTACCTGAGGGTCCCAATTTGGGCG 1860  
QY 1861 AAAATTCCTCACACAGATGGACACTTTTCAACCGTCTCTTATGGGCGGCTTTGGACTC 1920  
DB 1861 AAAATTCCTCACACAGATGGACACTTTTCAACCGTCTCTTATGGGCGGCTTTGGACTC 1920  
QY 1921 AAGAACCCGCTCTCAGATCCTCATAAAAACAACGCTGTTTCTTCCGAAATCTCTCGGCG 1980  
DB 1921 AAGAACCCGCTCTCAGATCCTCATAAAAACAACGCTGTTTCTTCCGAAATCTCTCGGCG 1980  
QY 1981 GAGTTTTCAGCTACAAAGTTTGTCTTATTCATCCCAATCTCAGGACAAAGTGA 2040  
DB 1981 GAGTTTTCAGCTACAAAGTTTGTCTTATTCATCCCAATCTCAGGACAAAGTGA 2040  
QY 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAG 2100  
DB 2041 GTGGAAATTTGAATGGAGCTGCAGAAAGAAACAGCAAGCGCTGGAATCCCGAAGTGCAG 2100  
QY 2101 TACACATCCAAATTTATGCAAAATCTGCCAAACGTTTGTATTTTACTGTGGACAAATGACAT 2160  
DB 2101 TACACATCCAAATTTATGCAAAATCTGCCAAACGTTTGTATTTTACTGTGGACAAATGACAT 2160  
QY 2161 TATATCTGAGCTCGGCCCAATTTGGGACCGCTTACCTTACCTTACCGTCCCTGT 2209  
DB 2161 TATATCTGAGCTCGGCCCAATTTGGGACCGCTTACCTTACCTTACCGTCCCTGT 2209

Search completed: November 29, 2005, 00:05:32  
Job time : 8078.36 secs

---

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 19:15:02 ; Search time 20.3389 Seconds  
(without alignments)  
12497.813 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_1\_143

Perfect score: 143

Sequence: 1 ttgcccactccctctctgcg.....aactccatcactagggttaa 143

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/1 COMB.seq.\*
- 2: /cgn2\_6/prodata/1/ina/5 COMB.seq.\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/1/ina/H COMB.seq.\*
- 6: /cgn2\_6/prodata/1/ina/PCUS COMB.seq.\*
- 7: /cgn2\_6/prodata/1/ina/PP COMB.seq.\*
- 8: /cgn2\_6/prodata/1/ina/RE COMB.seq.\*
- 9: /cgn2\_6/prodata/1/ina/backfilese1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	100.0	4718	3	US-09-807-802A-1
2	137.2	95.9	4693	3	US-09-807-802A-19
3	119	83.2	4718	3	US-09-807-802A-1
4	107.4	75.1	145	2	US-07-789-917A-1
5	107.4	75.1	145	3	US-08-702-573-4
6	107.4	75.1	145	3	US-07-982-193-1
7	107.4	75.1	145	3	US-09-782-378A-6
8	107.4	75.1	165	2	US-07-989-841A-1
9	107.4	75.1	165	2	US-08-440-738A-1
10	107.4	75.1	165	3	US-08-471-914-1
11	107.4	75.1	165	3	US-09-276-625-7
12	107.4	75.1	165	3	US-10-054-665A-7
13	107.4	75.1	165	3	US-09-782-378A-8
14	107.4	75.1	212	3	US-09-276-625-6
15	107.4	75.1	212	3	US-10-054-665A-6
16	107.4	75.1	4675	3	US-09-782-378A-1
17	107.4	75.1	4675	3	US-09-782-378A-2
18	107.4	75.1	4679	3	US-10-038-972A-12
19	107.4	75.1	4680	2	US-08-254-358-1
20	107.4	75.1	4680	2	US-08-475-391-1
21	107.4	75.1	4680	6	PCT-US95-07178-1
22	107.4	75.1	4681	3	US-09-807-802A-18
23	107.4	75.1	4681	3	US-09-807-802A-19
24	107.4	75.1	4693	3	US-09-807-802A-19

25	107.4	75.1	5932	3	US-09-299-141-4	Sequence 4, Appli
26	107.4	75.1	5932	3	US-09-299-141-4	Sequence 4, Appli
27	107.4	75.1	6142	3	US-09-299-141-8	Sequence 8, Appli
28	107.4	75.1	6142	3	US-09-299-141-8	Sequence 8, Appli
29	107.4	75.1	6253	3	US-08-893-327-15	Sequence 15, Appli
30	107.4	75.1	6253	3	US-08-893-327-15	Sequence 15, Appli
31	107.4	75.1	6280	3	US-08-893-327-17	Sequence 17, Appli
32	107.4	75.1	6280	3	US-08-893-327-17	Sequence 17, Appli
33	107.4	75.1	6280	3	US-08-893-327-19	Sequence 19, Appli
34	107.4	75.1	6280	3	US-08-893-327-19	Sequence 19, Appli
35	107.4	75.1	6565	3	US-09-299-141-1	Sequence 1, Appli
36	107.4	75.1	6565	3	US-09-299-141-1	Sequence 1, Appli
37	107.4	75.1	6714	3	US-09-299-141-6	Sequence 6, Appli
38	107.4	75.1	6714	3	US-09-299-141-6	Sequence 6, Appli
39	107.4	75.1	6924	3	US-09-299-141-9	Sequence 9, Appli
40	107.4	75.1	6924	3	US-09-299-141-9	Sequence 9, Appli
41	107.4	75.1	6924	3	US-09-299-141-10	Sequence 10, Appli
42	107.4	75.1	6924	3	US-09-299-141-10	Sequence 10, Appli
43	107.4	75.1	6924	3	US-09-299-141-11	Sequence 11, Appli
44	107.4	75.1	6924	3	US-09-299-141-11	Sequence 11, Appli
45	107.4	75.1	6981	3	US-09-299-141-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-807-802A-1  
; Sequence 1, Application US/09807802A  
; Patent No. 6759237  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,  
; TITLE OF INVENTION: Vectors and Host Cells Containing Same  
; FILE REFERENCE: GNPN.031USA  
; CURRENT APPLICATION NUMBER: US/09/807,802A  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/107,114  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: PCT/US99/25694  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 4718  
; TYPE: DNA  
; ORGANISM: AAV-1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (335)..(2206)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2223)..(4430)  
; OTHER INFORMATION:  
US-09-807-802A-1

Query Match 100.0%; Score 143; DB 3; Length 4718;

Best Local Similarity 100.0%; Pred. No. 2.7e-30;

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TTGCCACATCCCTCTCTCGCGCGCTCGCTCGGTGGGGCCCTGCGGACCAAGGTCCGC	60
Db	1	TTGCCACATCCCTCTCTCGCGCGCTCGCTCGGTGGGGCCCTGCGGACCAAGGTCCGC	60
QY	61	AGACGGCAGAGCTCTGCTCTTGGCGGCCCAACCGAGCGAGCGCGCAGAGGGAGTG	120
Db	61	AGACGGCAGAGCTCTGCTCTTGGCGGCCCAACCGAGCGAGCGCGCAGAGGGAGTG	120
QY	121	GGCACTCCATCACTAGGGGTAA	143
Db	121	GGCACTCCATCACTAGGGGTAA	143

```
RESULT 2
US-09-807-802A-19/c
; Sequence 19, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 4683
; TYPE: DNA
; ORGANISM: AAV-6
US-09-807-802A-19

Query Match      95.9%; Score 137.2; DB 3; Length 4683;
Best Local Similarity 97.9%; Pred. No. 1.1e-28;
Matches 139; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  TTGCCACCTCTCTCGCGGCTCGCTCGCTCGGTGGGGCTCGGGACCAAAGGTCCGC 60
DB      4683 TTGGCCACTCTCTATGCGCGCTGCTCGCTCGGTGGGGCTCGGGACCAAAGGTCCGC 4624

QY      61  AGACGGCAGAGCTCTGCTCGCGGCCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
DB      4623 AGACGGCAGAGCTCTGCTCGCGGCCGCCACCGAGCGAGCGCGCATAGAGGAGTG 4564

QY      121  GGCACTCCATCACTAGGGGTA 142
DB      4563 GGCACTCCATCACTAGGGGTA 4542

RESULT 3
US-09-807-802A-1/c
; Sequence 1, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4718
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (335)..(2206)
; OTHER INFORMATION:
; NAME/KEY: CDS
; LOCATION: (2223)..(4430)
; OTHER INFORMATION:
```

## US-09-807-802A-1

```
Query Match      83.2%; Score 119; DB 3; Length 4718;
Best Local Similarity 89.5%; Pred. No. 1.1e-23;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1  TTGCCACTCCCTCTCTCGCGGCTCGCTCGCTCGGTGGGGCTCGGGACCAAAGGTCCGC 60
DB      4718 TTGCCACTCCCTCTCTCGCGGCTCGCTCGCTCGGTGGGGCTCGGGACCAAAGGTCCGC 4659

QY      61  AGACGGCAGAGCTCTGCTCTGCGGCCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
DB      4658 CGTCTGGGGACCTTTGGTCCGACGCCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 4599

QY      121  GGCACTCCATCACTAGGGGTA 143
DB      4598 GGCACTCCATCACTAGGGGTA 4576

RESULT 4
US-07-789-917A-1
; Sequence 1, Application US/07789917A
; Patent No. 5252479
; GENERAL INFORMATION:
; APPLICANT: Srivastava, Arun
; TITLE OF INVENTION: SAFE VECTOR FOR GENE THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release q.0, Version q.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/789,917A
; FILING DATE: 19911118
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McNulty, William E.
; REGISTRATION NUMBER: 22,606
; REFERENCE/DOCKET NUMBER: 8361
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-789-917A-1

Query Match      75.1%; Score 107.4; DB 2; Length 145;
Best Local Similarity 85.1%; Pred. No. 1.2e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY      1  TTGCCACTCCCTCTCTCGCGGCTCGCTCGCTCGGTGGGGCTCGGGACCAAAGGTCCGC 60
DB      1  TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGGGCGGACCAAAGGTCCGC 60

QY      61  AGACGGCAGAGCTCTGCTCTGCGGCCGCCACCGAGCGAGCGCGCAGAGAGGAGTG 120
DB      61  CGACGCCCGGGCTTTGCCCGGCGCCTCAGTGTAGCGAGCGCGCAGAGAGGAGTG 120

QY      121  GGCACTCCATCACTAGGGGT 141
```

Db 121 GCCAACTCCATCACTAGGGT 141

RESULT 5

US-08-702-573-4

Sequence 4, Application US/08702573

Patent No. 6033885

GENERAL INFORMATION:

APPLICANT: LATTA, Martine

APPLICANT: DENEPIE, Patrice

APPLICANT: VIGNE, Emmanuelle

APPLICANT: PERRICAUDET, Michel

TITLE OF INVENTION: INTEGRATIVE RECOMBINANT ADENOVIRUSES, PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF

TITLE OF INVENTION: PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Inc.

STREET: 500 Arcola Rd. 3C43

CITY: Collegeville

STATE: PA

COUNTRY: USA

ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/702,573

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 94/02445

FILING DATE: 03-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/FR95/00233

FILING DATE: 28-FEB-1995

ATTORNEY/AGENT INFORMATION:

NAME: Smith Ph.D., Julie K.

REGISTRATION NUMBER: 38,619

REFERENCE/DOCKET NUMBER: ST94011-US

TELEPHONE: (610)454-3839

TELEFAX: (610)454-3808

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 145 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid

FEATURE:

NAME/KEY: misc feature

LOCATION: 1..145

OTHER INFORMATION: /note= "Minimal ITR Sequence"

US-08-702-573-4

Query Match 75.1%; Score 107.4; DB 3; Length 145;

Best Local Similarity 85.1%; Pred. No. 1.2e-20;

Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTGCGGACCAAAAGGTCCGC 60

Db 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTGCGGACCAAAAGGTCCGC 60

QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 120

Db 61 CGAGCGCGGGCTTTTGCCCGGGCCCTCAGTGAGCGAGCGCGCAGAGAGGAGTG 120

QY 121 GGCAACTCCATCACTAGGGT 141

Db 121 GGCAACTCCATCACTAGGGT 141

RESULT 6

US-07-982-193-1

Sequence 1, Application US/07982193

Patent No. 6261834

GENERAL INFORMATION:

APPLICANT: Srivastava, Arun

TITLE OF INVENTION: SAFE VECTOR FOR GENE THERAPY

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/982,193

FILING DATE: 19921125

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: McNulty, William E.

REGISTRATION NUMBER: 22,606

REFERENCE/DOCKET NUMBER: 8361

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 145 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-07-982-193-1

Query Match 75.1%; Score 107.4; DB 3; Length 145;

Best Local Similarity 85.1%; Pred. No. 1.2e-20;

Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTGCGGACCAAAAGGTCCGC 60

Db 1 TTGCCACTCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTGCGGACCAAAAGGTCCGC 60

QY 61 AGACGGCAGAGCTCTGCTCTGCGGCCCCACCGAGCGAGCGCGCAGAGAGGAGTG 120

Db 61 CGAGCGCGGGCTTTTGCCCGGGCCCTCAGTGAGCGAGCGCGCAGAGAGGAGTG 120

QY 121 GGCAACTCCATCACTAGGGT 141

Db 121 GGCAACTCCATCACTAGGGT 141

RESULT 7

US-09-782-378A-6

Sequence 6, Application US/09782378A

Patent No. 6916635

GENERAL INFORMATION:

APPLICANT: Hearing, Patrick

APPLICANT: Bahou, Wadie

APPLICANT: Sandalon, Ziv

APPLICANT: Gnatenko, Dmitri

TITLE OF INVENTION: Adenoviral Vectors

FILE REFERENCE: STONYB-04970

CURRENT APPLICATION NUMBER: US/09/782,378A

CURRENT FILING DATE: 2001-02-12



```
/ PRIOR APPLICATION NUMBER: 60/237,747
/ PRIOR FILING DATE: 2000-10-02
/ NUMBER OF SEQ ID NOS: 27
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 6
/ LENGTH: 145
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-782-378A-6

Query Match      75.1%; Score 107.4; DB 3; Length 145;
Best Local Similarity 85.1%; Pred. No. 1.2e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTCGGACCAAAAGGTCCGC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTCGGACCAAAAGGTCCGC 80
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGCGCGCAGAGAGGAGTG 120
Db 81 CGACGCCCGGGCTTTGCCCGGGCGGCTCTAGTGAGCGAGCGCGCGCAGAGAGGAGTG 140
QY 121 GGCAACTCCATCACTAGGGGT 141
Db 141 GCCAACTCCATCACTAGGGGT 161

RESULT 9
US-08-440-738A-1
/ Sequence 1, Application US/08440738A
/ Patent No. 5869305
/ GENERAL INFORMATION:
/ APPLICANT: Samuleki, R. J.
/ APPLICANT: Xiao, X.
/ TITLE OF INVENTION: Recombinant Viral Vector System
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Pennie & Edmonds
/ STREET: 1155 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036-2711
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/440,738A
/ FILING DATE: May 15, 1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Coruzzi, Laura A
/ REGISTRATION NUMBER: 30,742
/ REFERENCE/DOCKET NUMBER: 6636-022
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-8864/9741
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 165 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: unknown
/ MOLECULE TYPE: DNA (genomic)
US-08-440-738A-1

Query Match      75.1%; Score 107.4; DB 2; Length 165;
Best Local Similarity 85.1%; Pred. No. 1.2e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCTCGGACCAAAAGGTCCGC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCTCGGACCAAAAGGTCCGC 80
QY 61 AGACGGCAGAGCTCTGCTCTGCGCGCCGCCACCGAGCGAGCGCGCGCAGAGAGGAGTG 120
Db 81 CGACGCCCGGGCTTTGCCCGGGCGGCTCTAGTGAGCGAGCGCGCGCAGAGAGGAGTG 140
QY 121 GGCAACTCCATCACTAGGGGT 141
Db 141 GCCAACTCCATCACTAGGGGT 161

RESULT 10
US-07-989-841A-1
/ Sequence 1, Application US/07989841A
/ Patent No. 5478745
/ GENERAL INFORMATION:
/ APPLICANT: Samuleki, R. J.
/ APPLICANT: Xiao, X.
/ TITLE OF INVENTION: Recombinant Viral Vector System
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Pennie & Edmonds
/ STREET: 1155 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036-2711
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/989,841A
/ FILING DATE: On even date herewith
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Coruzzi, Laura A
/ REGISTRATION NUMBER: 30,742
/ REFERENCE/DOCKET NUMBER: 6636-013
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-8864/9741
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 165 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: unknown
/ MOLECULE TYPE: DNA (genomic)
US-07-989-841A-1

Query Match      75.1%; Score 107.4; DB 2; Length 165;
Best Local Similarity 85.1%; Pred. No. 1.2e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
```

<p>US-08-471-914-1</p> <p>; Sequence 1, Application US/08471914A ; Patent No. 6057152 ; GENERAL INFORMATION: ; APPLICANT: Samuleki, R. ; APPLICANT: Xiso, X. ; TITLE OF INVENTION: RECOMBINANT VIRAL VECTOR SYSTEM ; FILE REFERENCE: 6636-027 ; CURRENT APPLICATION NUMBER: US/08/471,914A ; CURRENT FILING DATE: 1995-06-06 ; EARLIER APPLICATION NUMBER: 08/440,738 ; EARLIER FILING DATE: 1995-05-15 ; NUMBER OF SEQ ID NOS: 13 ; SOFTWARE: Patentin Ver. 2.0 ; SEQ ID NO 1 ; LENGTH: 165 ; TYPE: DNA ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Description of Artificial Sequence: double-D US-08-471-914-1</p>	<p>Query Match            75.1%; Score 107.4; DB 3; Length 165; Best Local Similarity   85.1%; Pred. No. 1.2e-20; Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;</p> <p>Qy     1   TTGCCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCCTTCGCGACCAAAGGTCCGC 60 Db     21   TTGGCACTCCCTCTCTGCGCGCTCGCTCGTACTGAGCGGGCGGACCAAAGGTCCGC 80</p> <p>Qy     61   AGACGGCAGAGCTCTGCTCTTGCGCGCCCCACCAGCGAGCGCGCGCACGAGGAGTG 120 Db     81   CGAGCGCGGGCTTTGCGGGCGGCTCAGTGAGCGAGCGAGCGCGCGAGAGGAGTG 140</p> <p>Qy     121   GGCACTCCATCATCTAGGGGT 141 Db     141   GCCAACTCCATCATCTAGGGGT 161</p> <p>RESULT 11</p> <p>US-09-276-625-7 ; Sequence 7, Application US/09276625 ; Patent No. 6436392 ; GENERAL INFORMATION: ; APPLICANT: Engelhardt, John F. ; APPLICANT: Duan, Dongsheng ; TITLE OF INVENTION: Adeno-associated virus vectors ; FILE REFERENCE: 875.007U1 ; CURRENT APPLICATION NUMBER: US/09/276,625 ; CURRENT FILING DATE: 1999-03-25 ; PRIOR APPLICATION NUMBER: US 60/086,166 ; PRIOR FILING DATE: 1998-05-20 ; NUMBER OF SEQ ID NOS: 13 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 7 ; LENGTH: 165 ; TYPE: DNA ; ORGANISM: Unknown ; FEATURE: ; OTHER INFORMATION: sequence US-09-276-625-7</p> <p>Query Match            75.1%; Score 107.4; DB 3; Length 165; Best Local Similarity   85.1%; Pred. No. 1.2e-20; Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;</p> <p>Qy     1   TTGCCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCCTTCGCGACCAAAGGTCCGC 60 Db     21   TTGGCACTCCCTCTCTGCGCGCTCGCTCGTACTGAGCGGGCGGACCAAAGGTCCGC 80</p> <p>Qy     61   AGACGGCAGAGCTCTGCTCTTGCGCGCCCCACCAGCGAGCGCGCGCACGAGGAGTG 120 Db     81   CGAGCGCGGGCTTTGCGGGCGGCTCAGTGAGCGAGCGAGCGCGCGAGAGGAGTG 140</p> <p>Qy     121   GGCACTCCATCATCTAGGGGT 141 Db     141   GCCAACTCCATCATCTAGGGGT 161</p>
<p>US-08-471-914-2</p> <p>; Sequence 1, Application US/08471914A ; Patent No. 6057152 ; GENERAL INFORMATION: ; APPLICANT: Samuleki, R. ; APPLICANT: Xiso, X. ; TITLE OF INVENTION: RECOMBINANT VIRAL VECTOR SYSTEM ; FILE REFERENCE: 6636-027 ; CURRENT APPLICATION NUMBER: US/08/471,914A ; CURRENT FILING DATE: 1995-06-06 ; EARLIER APPLICATION NUMBER: 08/440,738 ; EARLIER FILING DATE: 1995-05-15 ; NUMBER OF SEQ ID NOS: 13 ; SOFTWARE: Patentin Ver. 2.0 ; SEQ ID NO 1 ; LENGTH: 165 ; TYPE: DNA ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Description of Artificial Sequence: double-D US-08-471-914-2</p>	<p>Query Match            75.1%; Score 107.4; DB 3; Length 165; Best Local Similarity   85.1%; Pred. No. 1.2e-20; Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;</p> <p>Qy     1   TTGCCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCCTTCGCGACCAAAGGTCCGC 60 Db     21   TTGGCACTCCCTCTCTGCGCGCTCGCTCGTACTGAGCGGGCGGACCAAAGGTCCGC 80</p> <p>Qy     61   AGACGGCAGAGCTCTGCTCTTGCGCGCCCCACCAGCGAGCGCGCGCACGAGGAGTG 120 Db     81   CGAGCGCGGGCTTTGCGGGCGGCTCAGTGAGCGAGCGAGCGCGCGAGAGGAGTG 140</p> <p>Qy     121   GGCACTCCATCATCTAGGGGT 141 Db     141   GCCAACTCCATCATCTAGGGGT 161</p> <p>RESULT 12</p> <p>US-10-054-665A-7 ; Sequence 7, Application US/10054665A ; Patent No. 6897045 ; GENERAL INFORMATION: ; APPLICANT: Engelhardt, John F. ; APPLICANT: Duan, Dongshen ; APPLICANT: University of Iowa Research Foundation ; TITLE OF INVENTION: Adeno-associated virus vectors ; FILE REFERENCE: 875.007U2 ; CURRENT APPLICATION NUMBER: US/10/054,665A ; CURRENT FILING DATE: 2002-01-22 ; PRIOR APPLICATION NUMBER: US 60/086,166 ; PRIOR FILING DATE: 1998-05-20 ; PRIOR APPLICATION NUMBER: US 09/276,625 ; PRIOR FILING DATE: 1999-03-25 ; NUMBER OF SEQ ID NOS: 14 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 7 ; LENGTH: 165 ; TYPE: DNA ; ORGANISM: Unknown ; FEATURE: ; OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. 6897045 5,478,745 US-10-054-665A-7</p>
<p>US-08-471-914-3</p> <p>; Sequence 1, Application US/08471914A ; Patent No. 6057152 ; GENERAL INFORMATION: ; APPLICANT: Samuleki, R. ; APPLICANT: Xiso, X. ; TITLE OF INVENTION: RECOMBINANT VIRAL VECTOR SYSTEM ; FILE REFERENCE: 6636-027 ; CURRENT APPLICATION NUMBER: US/08/471,914A ; CURRENT FILING DATE: 1995-06-06 ; EARLIER APPLICATION NUMBER: 08/440,738 ; EARLIER FILING DATE: 1995-05-15 ; NUMBER OF SEQ ID NOS: 13 ; SOFTWARE: Patentin Ver. 2.0 ; SEQ ID NO 1 ; LENGTH: 165 ; TYPE: DNA ; ORGANISM: Artificial Sequence ; FEATURE: ; OTHER INFORMATION: Description of Artificial Sequence: double-D US-08-471-914-3</p>	<p>Query Match            75.1%; Score 107.4; DB 3; Length 165; Best Local Similarity   85.1%; Pred. No. 1.2e-20; Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;</p> <p>Qy     1   TTGCCCACTCCCTCTCTGCGCGCTCGCTCGGTGGGGCCTTCGCGACCAAAGGTCCGC 60 Db     21   TTGGCACTCCCTCTCTGCGCGCTCGCTCGTACTGAGCGGGCGGACCAAAGGTCCGC 80</p> <p>Qy     61   AGACGGCAGAGCTCTGCTCTTGCGCGCCCCACCAGCGAGCGCGCGCACGAGGAGTG 120 Db     81   CGAGCGCGGGCTTTGCGGGCGGCTCAGTGAGCGAGCGAGCGCGCGAGAGGAGTG 140</p> <p>Qy     121   GGCACTCCATCATCTAGGGGT 141 Db     141   GCCAACTCCATCATCTAGGGGT 161</p> <p>RESULT 13</p> <p>US-09-782-378A-8 ; Sequence 8, Application US/09782378A ; Patent No. 6916635 ; GENERAL INFORMATION: ; APPLICANT: Hearing, Patrick ; APPLICANT: Bahou, Wadie ; APPLICANT: Sandalon, Ziv ; APPLICANT: Gnatenko, Dmitri ; TITLE OF INVENTION: Adenoviral Vectors ; FILE REFERENCE: STONYB-04970 ; CURRENT APPLICATION NUMBER: US/09/782,378A ; CURRENT FILING DATE: 2001-02-12 ; PRIOR APPLICATION NUMBER: 60/237,747 ; PRIOR FILING DATE: 2000-10-02 ; NUMBER OF SEQ ID NOS: 27 ; SOFTWARE: Patentin version 3.0 ; SEQ ID NO 8 ; LENGTH: 165 ; TYPE: DNA ; ORGANISM: Homo sapiens US-09-782-378A-8</p>

Query Match 75.1%; Score 107.4; DB 3; Length 165;  
Best Local Similarity 85.1%; Pred. No. 1.2e-20;  
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
QY 1 TTGCCCACCTCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCCAAAGGTCCGC 60  
DB 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCCGC 80  
QY 61 AGACGGCAGAGCTCTGCTTGGCGGCCACCACGAGGAGCGAGCGGCGAGAGAGGAGTG 120  
DB 81 CGACGCCCGGGCTTTGGCGCGCGCTCAGTGAGCGAGCGAGCGCGAGAGAGGAGTG 140  
QY 121 GGCAACTCCATCACTAGGGGT 141  
DB 141 GGCAACTCCATCACTAGGGGT 161

RESULT 14  
US-09-276-625-6  
; Sequence 6, Application US/09276625  
; Patent No. 6436392  
; GENERAL INFORMATION:  
; APPLICANT: Engelhardt, John F.  
; TITLE OF INVENTION: Adeno-associated virus vectors  
; FILE REFERENCE: 875.007US1  
; CURRENT APPLICATION NUMBER: US/09/276,625  
; PRIOR FILING DATE: 1999-03-25  
; PRIOR APPLICATION NUMBER: US 60/086,166  
; PRIOR FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; TYPE: DNA  
; ORGANISM: AAV circular intermediate, clone p1202  
US-09-276-625-6

Query Match 75.1%; Score 107.4; DB 3; Length 272;  
Best Local Similarity 85.1%; Pred. No. 1.2e-20;  
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
QY 1 TTGCCCACCTCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCCAAAGGTCCGC 60  
DB 69 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCCGC 128  
QY 61 AGACGGCAGAGCTCTCTGCGCGGCCACCACGAGCGAGCGCGAGAGAGGAGTG 120  
DB 129 CGACGCCCGGGCTTTGGTGGCGCGCTCAGTGAGCGAGCGAGCGCGAGAGGAGTG 188  
QY 121 GGCAACTCCATCACTAGGGGT 141  
DB 189 GGCAACTCCATCACTAGGGGT 209

RESULT 15  
US-10-054-665A-6  
; Sequence 6, Application US/10054665A  
; Patent No. 6897045  
; GENERAL INFORMATION:  
; APPLICANT: Engelhardt, John F.  
; TITLE OF INVENTION: University of Iowa Research Foundation  
; FILE REFERENCE: 875.007US2  
; CURRENT APPLICATION NUMBER: US/10/054,665A  
; PRIOR FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: US 60/086,166  
; PRIOR FILING DATE: 1998-05-20  
; PRIOR APPLICATION NUMBER: US 09/276,625  
; PRIOR FILING DATE: 1999-03-25  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6  
; LENGTH: 272  
; TYPE: DNA  
; ORGANISM: AAV circular intermediate, clone p1202  
US-10-054-665A-6  
Query Match 75.1%; Score 107.4; DB 3; Length 272;  
Best Local Similarity 85.1%; Pred. No. 1.2e-20;  
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
QY 1 TTGCCCACCTCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCTTGGGACCCAAAGGTCCGC 60  
DB 69 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCCGC 128  
QY 61 AGACGGCAGAGCTCTGCTTGGCGGCCACCACGAGCGAGCGCGAGAGAGGAGTG 120  
DB 129 CGACGCCCGGGCTTTGGTGGCGCGCTCAGTGAGCGAGCGAGCGCGAGAGGAGTG 188  
QY 121 GGCAACTCCATCACTAGGGGT 141  
DB 189 GGCAACTCCATCACTAGGGGT 209

Search completed: November 29, 2005, 14:27:39  
Job time : 21.3389 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2005, 19:15:02 ; Search time 20.3389 Seconds  
(without alignments)  
12497.813 Million cell updates/sec

Title: US-10-696-900-1\_COPY\_4576\_4718

Perfect score: 143

Sequence: 1 ttacccttagtgatggagtt.....cgccagagggagtgaggca 143

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:\*
- 1: /cgn2\_6/prodata/1/ina/1 COMB.seq:\*
  - 2: /cgn2\_6/prodata/1/ina/5 COMB.seq:\*
  - 3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*
  - 4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*
  - 5: /cgn2\_6/prodata/1/ina/H COMB.seq:\*
  - 6: /cgn2\_6/prodata/1/ina/PCrus COMB.seq:\*
  - 7: /cgn2\_6/prodata/1/ina/PP COMB.seq:\*
  - 8: /cgn2\_6/prodata/1/ina/RE COMB.seq:\*
  - 9: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143	100.0	4718	3	US-09-807-802A-1
2	120.2	84.1	4683	3	US-09-807-802A-19
3	119	83.2	4718	3	US-09-807-802A-1
4	113.2	79.2	4683	3	US-09-807-802A-19
5	109	76.2	4681	3	US-09-807-802A-18
6	107.4	75.1	145	3	US-08-525-866-1
7	107.4	75.1	165	2	US-07-989-841A-1
8	107.4	75.1	165	2	US-08-440-738A-1
9	107.4	75.1	165	3	US-08-471-914-1
10	107.4	75.1	165	3	US-09-276-625-7
11	107.4	75.1	165	3	US-10-054-665A-7
12	107.4	75.1	192	3	US-08-702-573-3
13	107.4	75.1	272	3	US-09-276-625-6
14	107.4	75.1	272	3	US-10-054-665A-6
15	107.4	75.1	4675	3	US-09-782-378A-1
16	107.4	75.1	4675	3	US-09-782-378A-2
17	107.4	75.1	4679	3	US-10-038-972A-12
18	107.4	75.1	4680	2	US-08-254-358-1
19	107.4	75.1	4680	2	US-08-475-391-1
20	107.4	75.1	4680	2	US-08-709-609-1
21	107.4	75.1	4680	6	PCT-US95-07178-1
22	107.4	75.1	8698	3	US-09-770-315-2
23	105.8	74.0	272	3	US-09-276-625-4
24	105.8	74.0	272	3	US-10-054-665A-4

c 25	105.4	73.7	174	3	US-09-394-110A-1	Sequence 1, Appli
c 26	102.6	71.7	272	3	US-09-276-625-6	Sequence 6, Appli
c 27	102.6	71.7	272	3	US-10-054-665A-6	Sequence 6, Appli
c 28	101	70.6	183	3	US-09-394-110A-2	Sequence 2, Appli
c 29	99.4	69.5	145	2	US-07-789-917A-1	Sequence 1, Appli
c 30	99.4	69.5	145	3	US-08-702-573-4	Sequence 4, Appli
c 31	99.4	69.5	145	3	US-07-982-193-1	Sequence 1, Appli
c 32	99.4	69.5	145	3	US-09-782-378A-6	Sequence 6, Appli
c 33	99.4	69.5	145	2	US-07-989-841A-1	Sequence 1, Appli
c 34	99.4	69.5	165	2	US-08-440-738A-1	Sequence 1, Appli
c 35	99.4	69.5	165	3	US-08-471-914-1	Sequence 7, Appli
c 36	99.4	69.5	165	3	US-09-276-625-7	Sequence 7, Appli
c 37	99.4	69.5	165	3	US-10-054-665A-7	Sequence 7, Appli
c 38	99.4	69.5	165	3	US-09-782-378A-8	Sequence 8, Appli
c 39	99.4	69.5	165	3	US-09-782-378A-1	Sequence 2, Appli
c 40	99.4	69.5	4675	3	US-09-782-378A-2	Sequence 2, Appli
c 41	99.4	69.5	4679	3	US-10-038-972A-12	Sequence 12, Appli
c 42	99.4	69.5	4680	2	US-08-254-358-1	Sequence 1, Appli
c 43	99.4	69.5	4680	2	US-08-475-391-1	Sequence 1, Appli
c 44	99.4	69.5	4680	2	US-08-709-609-1	Sequence 1, Appli
c 45	99.4	69.5	4680	6	PCT-US95-07178-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-807-802A-1  
; Sequence 1, Application US/09807802A  
; Patent No. 6759237  
; GENERAL INFORMATION:  
; APPLICANT: Wilson, James M.  
; APPLICANT: Xiao, Weidong  
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,  
; TITLE OF INVENTION: Vectors and Host Cells Containing Same  
; FILE REFERENCE: GNPVN.031USA  
; CURRENT APPLICATION NUMBER: US/09/807,802A  
; PRIOR FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: US 60/107,114  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: PCT/US99/25694  
; PRIOR FILING DATE: 1999-11-02  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 4718  
; TYPE: DNA  
; ORGANISM: AAV-1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (335)..(2206)  
; OTHER INFORMATION:  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2223)..(4430)  
; OTHER INFORMATION:  
US-09-807-802A-1

Query Match	100.0%	Score 143;	DB 3;	Length 4718;
Best Local Similarity	100.0%	Pred. No. 2.2e-30;		
Matches 143;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TTACCCCTAGTGTGAGTTGCCACTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCC	60	
Db	4576	TTACCCCTAGTGTGAGTTGCCACTCCCTCTCTCGCGCGCTCGCTCGCTCGGTGGGGCC	4635	
QY	61	TGCGGACCAAGGTCCGACAGCGGAGCTTGTCTTGCCTGCCGCCACCGACGACGCA	120	
Db	4636	TGCGGACCAAGGTCCGACAGCGGAGCTTGTCTTGCCTGCCGCCACCGACGACGCA	4695	
QY	121	GCGCGCAGAGAGGAGTGGGCAA	143	
Db	4696	GCGCGCAGAGAGGAGTGGGCAA	4718	

```
RESULT 2
US-09-807-802A-19/c
; Sequence 19, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 4683
; TYPE: DNA
; ORGANISM: AAV-6
US-09-807-802A-19

Query Match      84.1%; Score 120.2; DB 3; Length 4683;
Best Local Similarity 97.6%; Pred. No. 4.3e-24;
Matches 122; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 19 TTGCCACATCCCTCTCTGCGCGCTCGCTCGCTGGGGGCTTGGCGACCAAGGTCGCG 78
DB 4683 TTGGCCACATCCCTCTATGCGCTCGCTCGCTGGGGGCTTGGCGACCAAGGTCGCG 4624

QY 79 AGACGGCAGAGCTCTGCTCTGCGGCCCCCAGCGAGCGCGCAGAGGGAGTG 138
DB 4623 AGACGGCAGAGCTCTGCTCTGCGGCCCCCAGCGAGCGCGCAGAGGGAGTG 4564

QY 139 GGCAG 143
DB 4563 GGCAG 4559

RESULT 3
US-09-807-802A-1/c
; Sequence 1, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4718
; TYPE: DNA
; ORGANISM: AAV-1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (335)..(2206)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2223)..(4430)
; OTHER INFORMATION:
```

```
US-09-807-802A-1
Query Match      83.2%; Score 119; DB 3; Length 4718;
Best Local Similarity 89.5%; Pred. No. 9.1e-24;
Matches 128; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 TTACCCCTAGTAGTGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCC 60
DB 143 TTACCCCTAGTAGTGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCC 84

QY 61 TGGCGACCAAGGTCGCGAGCGGAGAGCTCTGCTCTGCGGCCCCCAGCGAGCGAG 120
DB 83 GCGCAGCAGAGCTCTGCGGCTGCGGACCTTTGGTCCGAGGCCCCACCGAGCGAGCGA 24

QY 121 GCGCGCAGAGGGAGTGCGCAA 143
DB 23 GCGCGCAGAGGGAGTGCGCAA 1

RESULT 4
US-09-807-802A-19
; Sequence 19, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 4683
; TYPE: DNA
; ORGANISM: AAV-6
US-09-807-802A-19

Query Match      79.2%; Score 113.2; DB 3; Length 4683;
Best Local Similarity 87.3%; Pred. No. 3.6e-22;
Matches 124; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 2 TACCCCTAGTAGTGAGTTGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGTGGGGCC 61
DB 4542 TACCCCTAGTAGTGAGTTGCCACTCCCTCTATGCGGCTCGCTCGCTCGGTGGGGCC 4601

QY 62 GCGGACCAAGGTCGCGAGCGGAGAGCTCTGCTCTGCGGCCCCCAGCGAGCGAG 121
DB 4602 GCAGAGCAGAGCTCTGCGGCTCTGCGGACCTTTGGTCCGAGGCCCCACCGAGCGAG 4661

QY 122 GCGCGCAGAGGGAGTGCGCAA 143
DB 4662 GCGCGCAGAGGGAGTGCGCAA 4683

RESULT 5
US-09-807-802A-18
; Sequence 18, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNPVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
```

```
/ PRIOR FILING DATE: 1998-11-05
/ PRIOR APPLICATION NUMBER: PCT/US99/25694
/ PRIOR FILING DATE: 1999-11-02
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 18
/ LENGTH: 4681
/ TYPE: DNA
/ ORGANISM: AAV-2
US-09-807-802A-18

Query Match          76.2%; Score 109; DB 3; Length 4681;
Best Local Similarity 85.8%; Pred. No. 5.2e-21;
Matches 121; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 3 ACCCTAGTGATGGAGTTGCCACATCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 62
Db 141 ACCCTAGTGATGGAGTTGCCACATCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 82

QY 63 CGGACCAAGGTCGCCAGACGCGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGAGC 122
Db 81 GCGACCAAGGTCGCCAGACGCGCGGCTTTGCCGGGGGCTCAGTGAGCGAGCGAGC 22

QY 123 GCGCAGAGAGGGAGTGGGCAA 143
Db 21 GCGCAGAGAGGGAGTGGGCAA 1

RESULT 7
US-07-989-841A-1
/ Sequence 1, Application US/07989841A
/ Patent No. 5478745
/ GENERAL INFORMATION:
/ APPLICANT: Samuleki, R. J.
/ APPLICANT: Xiao, X.
/ TITLE OF INVENTION: Recombinant Viral Vector System
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Pennie & Edmonds
/ STREET: 1155 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036-2711
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/989,841A
/ FILING DATE: On even date herewith
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Coruzzi, Laura A
/ REGISTRATION NUMBER: 30,742
/ REFERENCE/DOCKET NUMBER: 6636-013
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-8864/9741
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 165 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: unknown
/ MOLECULE TYPE: DNA (genomic)
US-07-989-841A-1

Query Match          75.1%; Score 107.4; DB 2; Length 165;
Best Local Similarity 85.1%; Pred. No. 1e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 ACCCTAGTGATGGAGTTGCCACATCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 62
Db 5 ACCCTAGTGATGGAGTTGCCACATCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 64

QY 63 CGGACCAAGGTCGCCAGACGCGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGAGC 122
Db 65 GCGACCAAGGTCGCCAGACGCGCGGCTTTGCCGGGGGCTCAGTGAGCGAGCGAGC 124

QY 123 GCGCAGAGAGGGAGTGGGCAA 143
Db 125 GCGCAGAGAGGGAGTGGGCAA 145

RESULT 8

/ PRIOR FILING DATE: 1998-11-05
/ PRIOR APPLICATION NUMBER: PCT/US99/25694
/ PRIOR FILING DATE: 1999-11-02
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 18
/ LENGTH: 4681
/ TYPE: DNA
/ ORGANISM: AAV-2
US-09-807-802A-18

Query Match          76.2%; Score 109; DB 3; Length 4681;
Best Local Similarity 85.8%; Pred. No. 5.2e-21;
Matches 121; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 3 ACCCTAGTGATGGAGTTGCCACATCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 62
Db 4541 ACCCTAGTGATGGAGTTGCCACATCCCTCTCTGCGCGCTCGCTCGTGGGGCCCTG 4600

QY 63 CGGACCAAGGTCGCCAGACGCGAGCTCTGCTCTGCCGGCCCCACCGAGCGAGCGAGC 122
Db 4601 GCGACCAAGGTCGCCAGACGCGCGGCTTTGCCGGGGGCTCAGTGAGCGAGCGAGC 4660

QY 123 GCGCAGAGAGGGAGTGGGCAA 143
Db 4661 GCGCAGAGAGGGAGTGGGCAA 4681

RESULT 6
US-08-525-866-1/c
/ Sequence 1, Application US/08525866
/ Patent No. 6207457
/ GENERAL INFORMATION:
/ APPLICANT: NATSOUIS, GEORGES
/ APPLICANT: FURSKY, RICHARD T.
/ TITLE OF INVENTION: TARGETED NUCLEOTIDE SEQUENCE DELIVERY
/ TITLE OF INVENTION: AND INTEGRATION SYSTEM
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: REED & ROBINS
/ STREET: 285 Hamilton Avenue, Suite 200
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94301
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/525,866
/ FILING DATE: 08-SEP-1995
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: ROBINS, ROBERTA L.
/ REGISTRATION NUMBER: 33,208
/ REFERENCE/DOCKET NUMBER: 0800-0006
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 327-3400
/ TELEFAX: (415) 327-3231
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 145 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
US-08-525-866-1

Query Match          75.1%; Score 107.4; DB 3; Length 145;
Best Local Similarity 85.1%; Pred. No. 9.9e-21;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
```



```
US-08-440-738A-1
; Sequence 1, Application US/08440738A
; Patent No. 5869305
; GENERAL INFORMATION:
; APPLICANT: Samuleki, R. J.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: Recombinant Viral Vector System
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,738A
; FILING DATE: May 15, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6636-022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8064/9741
; TLEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-440-738A-1

Query Match 75.1%; Score 107.4; DB 2; Length 165;
Best Local Similarity 85.1%; Pred. No. 1e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 ACCCTAGTGGAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCGCTG 62
Db 5 ACCCTAGTGGAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCGCTG 64

QY 63 CGGACCAAGGTCCGACAGCGGAGAGCTCTGCTCTGCGCGCGCGCGCGCGCGAGC 122
Db 65 GCGACCAAGGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGC 124

QY 123 GCGCAGAGAGGAGTGGGCAA 143
Db 125 GCGCAGAGAGGAGTGGGCAA 145

RESULT 9
US-08-471-914-1
; Sequence 1, Application US/08471914A
; Patent No. 6057152
; GENERAL INFORMATION:
; APPLICANT: Samuleki, R.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: RECOMBINANT VIRAL VECTOR SYSTEM
; CURRENT APPLICATION NUMBER: US/08/471,914A
; CURRENT FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/440,738
; EARLIER FILING DATE: 1995-05-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0

Query Match 75.1%; Score 107.4; DB 2; Length 165;
Best Local Similarity 85.1%; Pred. No. 1e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 ACCCTAGTGGAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCGCTG 62
Db 5 ACCCTAGTGGAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCGCTG 64

QY 63 CGGACCAAGGTCCGACAGCGGAGAGCTCTGCTCTGCGCGCGCGCGCGCGCGAGC 122
Db 65 GCGACCAAGGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGC 124

QY 123 GCGCAGAGAGGAGTGGGCAA 143
Db 125 GCGCAGAGAGGAGTGGGCAA 145

RESULT 10
US-09-276-625-7
; Sequence 7, Application US/09276625
; Patent No. 6436392
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongsheng
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US1
; CURRENT APPLICATION NUMBER: US/09/276,625
; CURRENT FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. 6436392 5,478,745
US-09-276-625-7

Query Match 75.1%; Score 107.4; DB 3; Length 165;
Best Local Similarity 85.1%; Pred. No. 1e-20;
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 ACCCTAGTGGAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCGCTG 62
Db 5 ACCCTAGTGGAGTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCGGCGCGCTG 64

QY 63 CGGACCAAGGTCCGACAGCGGAGAGCTCTGCTCTGCGCGCGCGCGCGCGCGAGC 122
Db 65 GCGACCAAGGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGC 124

QY 123 GCGCAGAGAGGAGTGGGCAA 143
Db 125 GCGCAGAGAGGAGTGGGCAA 145

RESULT 11
US-10-054-665A-7
; Sequence 7, Application US/10054665A
; Patent No. 6897045
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongsheng
```



QY 63 CGGACCAAGGTCGCCAGACGGCAGAGCTCTGCTCGCGGCCCCACCGAGCGAGCGAGC 122  
Db 4595 GCGACCAAAAGGTGCGCCCGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGAGC 4654  
QY 123 GCGCAGAGAGGGAGTGCGCAA 143  
Db 4655 GCGCAGAGAGGGAGTGCGCAA 4675

Search completed: November 29, 2005, 14:27:39  
Job time : 20.3389 secs

RESULT 14  
US-10-054-665A-6  
; Sequence 6, Application US/10054665A  
; Patent No. 6897045  
; GENERAL INFORMATION:  
; APPLICANT: Engelhardt, John F.  
; APPLICANT: Duan, Dongshen  
; APPLICANT: University of Iowa Research Foundation  
; TITLE OF INVENTION: Adeno-associated virus vectors  
; FILE REFERENCE: 875.007US2  
; CURRENT APPLICATION NUMBER: US/10/054,665A  
; CURRENT FILING DATE: 2002-01-22  
; PRIOR APPLICATION NUMBER: US 60/086,166  
; PRIOR FILING DATE: 1998-05-20  
; PRIOR APPLICATION NUMBER: US 09/276,625  
; PRIOR FILING DATE: 1999-03-25  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 272  
; TYPE: DNA  
; ORGANISM: AAV circular intermediate, clone p1202  
US-10-054-665A-6  
  
Query Match 75.1%; Score 107.4; DB 3; Length 272;  
Best Local Similarity 85.1%; Pred. No. 1.1e-20;  
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
  
QY 3 ACCCTAGTAGTGAGTGGCTCCCTCCCTCTCTGCGCGCTCGCTCGGTGGGCGCTG 62  
Db 53 ACCCTAGTAGTGAGTGGCTCCCTCTCTGCGCGCTCGCTCGGTGGGCGCTG 112  
  
QY 63 CGGACCAAGGTCGCCAGACGGCAGAGCTCTGCTCGCGGCCCCACCGAGCGAGCGAGC 122  
Db 113 GCGACCAAAAGGTGCGCCCGACGCCCGGGCTTTGGTCGCCCGGCTCAGTGAGCGAGCGAGC 172  
  
QY 123 GCGCAGAGAGGGAGTGCGCAA 143  
Db 173 GCGCAGAGAGGGAGTGCGCAA 193

RESULT 15  
US-09-782-378A-1  
; Sequence 1, Application US/09782378A  
; Patent No. 6916635  
; GENERAL INFORMATION:  
; APPLICANT: Hearing, Patrick  
; APPLICANT: Bahou, Wadie  
; APPLICANT: Sandaion, Ziv  
; APPLICANT: Gnatenko, Dmitri  
; TITLE OF INVENTION: Adenoviral Vectors  
; FILE REFERENCE: STONYB-04970  
; CURRENT APPLICATION NUMBER: US/09/782,378A  
; CURRENT FILING DATE: 2001-02-12  
; PRIOR APPLICATION NUMBER: 60/237,747  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 4675  
; TYPE: DNA  
; ORGANISM: Human adeno-associated virus 2  
US-09-782-378A-1  
  
Query Match 75.1%; Score 107.4; DB 3; Length 4675;  
Best Local Similarity 85.1%; Pred. No. 1.4e-20;  
Matches 120; Conservative 0; Mismatches 21; Indels 0; Gaps 0;  
  
QY 3 ACCCTAGTAGTGAGTGGCTCCCTCCCTCTCTGCGCGCTCGCTCGGTGGGCGCTG 62  
Db 4535 ACCCTAGTAGTGAGTGGCTCCCTCCCTCTCTGCGCGCTCGCTCGGTGGGCGCTG 4594